Measuring the Market Risk of Freight Rates: A Forecast Combination Approach

Christos Argyropoulos
University of Kent, UK

Ekaterini Panopoulou*
University of Kent, UK

June 18, 2017

Abstract

This paper addresses the issue of freight rate risk measurement via Value at Risk (VaR) and forecast combination methodologies while focusing on detailed performance evaluation. We contribute to the literature in three ways: First, we re-evaluate the performance of popular VaR estimation methods on freight rates amid the adverse economic consequences of the recent financial and sovereign debt crisis. Second, we provide a detailed and extensive backtesting and evaluation methodology. Last, we propose a forecast combination approach for estimating VaR. Our findings suggest that our combination methods produce more accurate estimates for all the sectors under scrutiny, while in some cases they may be viewed as conservative since they tend to overestimate nominal VaR.

Keywords: Backtesting; Combination Forecasts; Volatility Forecasts; Freight Rates; Performance Evaluation; Value-at-Risk;

Acknowledgements: This work was supported by the Economic and Social Research Council [grant number 1494761].

* *Corresponding author.* Ekaterini Panopoulou, Kent Business School, University of Kent, Canterbury CT2 7PE, United Kingdom, Tel.: 0044 1227 824469, Email: A.Panopoulou@kent.ac.uk.

1 Introduction

Freight rate risk has always been one of the most important risk factors of the shipping industry mainly because it affects its primary source of income. The uprising interest of participants in the shipping industry such as shipping companies, shipping hedge funds and shipping banks makes the accurate measurement of freight risk a procedure of high importance and difficulty induced by the intrinsic characteristics of the shipping industry itself. Specifically, the cyclicality of the maritime economy and the mechanics of the shipping markets create a complex profile for both the freight rates and their volatility.[1] While, from a financial perspective, freight rates are typically considered part of the commodities market, freight rate markets are quite different from the majority of the other commodities markets. For example, in contrast to all major traded commodities, freight rates are essentially not storable, a property that makes simple cost-of-carry valuations of futures contracts for freight rates impossible. Moreover, the freight rate spot market shows a high degree of volatility and seasonality which causes significant risks for shipowners, charterers and market participants in general (Alizadeh and Nomikos, 2011, Alizadeh *et al.*, 2014). Consequently, producing accurate estimates of freight risk is essential to freight market participants as it enables them to enhance their ability of sound strategic, investment and hedging decisions. This paper addresses the subject of freight rate risk measurement via Value at Risk (VaR) and forecast combination methodologies while focusing on detailed performance evaluation.

There are many factors behind the fluctuation of freight rates. In the long run, freight rates are determined through the interaction between the supply and demand schedules for shipping services. Information on vessel availability and production directly influences price levels. On the other hand, demand for shipping services is closely linked to the business cycle and economic growth through various channels. For example, during economic booms, both production of commodities and demand for crude oil is high. This leads to increases in both dry bulk (responsible for raw non liquid materials transportation) and tanker (responsible for the transportation of liquids, such as crude oil, petroleum products, etc.) freight rates. In the short run, the cost of operating a vessel greatly fluctuates with the cost of crude oil, mainly used as a shipping fuel and known as 'bunker'. Consequently, bunker prices are closely linked to crude oil prices and, therefore, freight rates react to changes in oil price levels. Moreover, tanker vessels account for approximately half of the world's seaborne trade and tanker freight rates determine the transportation cost of strategic for the world economy products such as the crude oil and its by-products. Finally, the importance of freight rates for global real economic activity is highlighted in the index of real economic activity constructed by Kilian (2009). This index is essentially based on global dry cargo freight rates and exhibits high correlation with the Bulk Dry Index (BDI). BDI is often employed as a general market indicator, "barometer", reflecting changes not only in the dry-bulk market, but in the overall worldwide real economy.

Given the importance of freight rates, calculating accurately freight rate risk is of utmost

[1] For an in depth discussion on freight rate characteristics, see Stopford (2009) and Alizadeh and Nomikos (2009).

importance for at least three reasons. First, market participants can develop hedging schemes more effectively and efficiently when they are aware of the risk they are exposed to. Simple risk metrics, like volatility, have not been proven adequate for this market due to deviation from normality, complexity, cyclicality and the existence of jumps during extreme events (see Kavussanos and Visvikis, 2004, Kavussanos and Dimitrakopoulos, 2011). For example, Alizadeh and Nomikos (2011) argue that volatility dynamics vary with shipping market conditions, i.e. they are regime dependent. In the same vein, Alizadeh *et al.* (2014) implement a regime switching multivariate approach in order to capture the volatility dynamics and possible correlations of spot and futures prices in the tanker sector. Second, during the last decade, the shipping freight market has transformed from a service market, where freight rate was the cost of transporting raw materials by sea to a market where freight rate is seen as an investment like any other asset or commodity (see e.g. Nomikos *et al.*, 2013). Market participants now include investment banks and hedge funds, who are interested in quantifying the risk profile of this alternative asset class having realized its potential benefit for both speculation and diversification. Alizadeh (2013) documents a positive contemporaneous relationship between trading volume, which has increased rapidly in the recent years, and volatility in the shipping forward freight market. Finally, VaR provides a mean of setting margin requirements in the freight exchange derivatives market, which is expanding fast. Given that forward freight agreements (FFAs) and freight options are employed to hedge freight rate risk and that trading of these derivatives can be done through an organized exchange, margin determination is very important. With the elimination of credit risk, margins reflect market liquidity and volatility of the underlying spot freight rates.[2] Specifically, FFAs are the primary instrument shipping market participants employ to hedge freight exposure risk. These contracts are agreements between a buyer and seller to settle the difference between the contract price and an appropriate settlement price in cash. The settlement price is normally the average of the spot freight rates on the underlying shipping route over the calendar month, reported by the Baltic Exchange. Alternative freight risk management techniques include time-charter contracts, contracts of affreightment (CoAs), and freight options. While period charter contracts and CoAs are considered physical forms of hedging, these contracts are not very liquid and operationally flexible. Finally, freight rate options can be used for hedging freight rates, but they are not very liquid and are comparatively expensive.

Despite the plethora of related research in the financial sector, the measurement of the market risk of freight rates has been under-researched. To the best of our knowledge the most recent contributions in the field of freight rate risk adopt the VaR methodology in order to measure the market risk of the dry bulk and tanker freight market. Specifically, Angelidis and

[2]Predictability in the underlying freight spot rate does not imply predictability of the corresponding derivative contract, since the standard cost-of-carry relationship for financial futures does not hold for the freight ones. This is because the underlying asset is not tradable, and hence the pricing by arbitrage argument cannot be applied. A series of papers have tested the unbiasedness hypothesis of the market, i.e. whether the freight rate market is efficient. Results are mixed mainly due to market segmentation (see Kavussanos and Visvikis, 2006, for a review). A more recent contribution by Goulas and Skiadopoulos (2012) points to the inefficiency of the International Maritime Exchange (IMAREX) freight futures markets over the daily horizon. The futures trading strategies based on the formed daily forecasts the authors develop yield a positive, economically significant risk premium.

Skiadopoulos (2008) applied a variety of parametric, non-parametric and hybrid methods in order to measure the market risk mainly for the dry bulk sector. Their findings suggest that in almost all cases the simplest non-parametric methods produce accurate results. On the other hand, Kavussanos and Dimitrakopoulos (2011) dealt with the selection of the appropriate freight rate risk model by applying a similar VaR methodology solely for the tanker sector. The authors find that parametric methods are more suitable for this sector. More recently, Abouarghoub *et al.*(2014) utilized a two state Markov-Switching distinctive conditional variance model in order to improve the tanker sector VaR forecasts. Their results suggest that a regime switching approach can capture more precisely the tanker sector volatility dynamics, thus providing better VaR forecasts.

In this paper, we contribute to the literature on both tanker and dry bulk freight rate risk forecasting in the following dimensions. First, we re-evaluate the performance of popular VaR estimation methods amid the adverse economic consequences of the recent financial and sovereign debt crisis. Second, we provide a detailed and extensive backtesting methodology in order to identify possible weaknesses associated with the standard backtesting criteria. Finally, we propose a forecast combination approach in estimating VaR, which provides more accurate VaR estimates while reducing the cost in time and resources.

More specifically, we calculate the daily 5% and 1% VaR of a long position comprising the most important Baltic Exchange indices and individual routes by applying a wide range of estimation methods. The aggregate freight rate indices employed are averages of individual route or time chartered indices and can be thought of as portfolios of freight rate positions covering large fleets of vessels. On the other hand, individual route indices are more relevant for market risk exposures of smaller companies, since these companies employ vessels in one of these routes which also serve as the underlying assets of freight rate derivatives (see Kavussanos and Dimitrakopoulos, 2011). To account for both types of risk exposure, we complement our analysis with the most actively utilized individual routes of both the dry bulk and the tanker sector. Detailing our methodology, we apply parametric, non-parametric, hybrid and a variety of combination methods on the logarithmic returns of the indices at hand and employ an evaluation sample that includes both booming periods and the crises periods of 2007-2009. Given the plethora of VaR estimation methods, we go one step further and investigate whether combinations of VaR forecasts can lead to gains such as: diversification gains, robustness to model mis-specification and structural breaks and bias correction of individual VaRs (see e.g. Timmermann, 2006 and Halbleib and Pohlmeier, 2012). To this end, we employ the mean, median and trimmed means (two versions) of individual VaR estimates aiming at producing a superior forecast. In order to conduct a more reliable and in depth evaluation, we implement a battery of newly developed backtesting criteria, namely the Engle and Manganelli (2004) quantile regression approach, the Christoffersen and Pelletier (2004) duration approach and the Colletaz *et al.* (2013) test along with the standard Christoffersen (1998) approach. More importantly, we complement statistical evaluation with the performance evaluation methodology proposed by Sener *et al.* (2012) in order to rank the implemented methods.

We find evidence that combination methods outperform the individual ones. Specifically we find that at the 5% VaR level, combination methods provide a globally accepted method which in every case examined produces equal or superior results to the highest ranked individual methods. Quite importantly, our findings are more pronounced at the 1% VaR level as we find that in the majority of cases combination methods are superior to the individual ones. We expect the empirical findings of the paper to be useful to a wider range of market participants given that the freight rate markets exhibit similar characteristics to financial markets, such as stock, bond, energy and commodity markets.

The remainder of this paper is organized as follows; Sections 2 and 3 present the theoretical framework and methodologies for computing and evaluating the VaR estimates. Section 4 presents the data and the empirical findings while Section 5 presents the related findings for individual route indices. Section 6 concludes the paper.

2 VaR Forecasting

There are three basic categories of risk measures; sensitivity measures, volatility measures and finally downside risk measures.[3] Downside risk measures are thought to be the most comprehensive category as they combine both sensitivity and volatility measures with the adverse effect of uncertainty. A typical example of a downside risk measure is Value at Risk (VaR) which can be described as the maximum potential deviation for a given significance level over a given time horizon. More specifically, let p_t be the price of the asset at the end of the day t and $r_t = \log\left(\frac{p_t}{p_{t-1}}\right)$ the daily log returns. For a long position, VaR is defined as the expected maximum loss over a specific horizon at a certain confidence level. For the continuous distribution case, it can be defined at a $q\%$ confidence level as follows:

$$P(r_t \leq -VaR\,(q)) = q \tag{1}$$

Equation (1) essentially defines VaR as the q quantile of the distribution of returns. Therefore, approximating the distribution of returns or more precisely its tails is essential to forecast VaR. In the following sections we describe the individual and combination VaR forecasting methods.

2.1 Individual Forecasting Methods

Approximating the returns distribution can be a cumbersome task given that the real Data Generating Process (DGP) is unknown. There is an abundance of proposed methods, which utilize the information set in an attempt to approximate accurately the unknown DGP. Based on the nature of this approximation, the corresponding methods are defined as parametric, non parametric and hybrid.

Despite the methods' specifications for the DGP approximation, individual forecasting methods may induce significant model risk to the forecasts. For instance, simple methods such as

[3]For an elaborate discussion, see Bessis (2002).

Historical Simulation (HS) or an unconditional distribution fit assume independent and identically distributed (iid) returns. This assumption simplifies significantly the approximation of the DGP as it provides an average on all possible events in the sample. However, assuming iid is counterintuitive given that empirical evidence (heteroscedasticity) point towards returns dependency. In addition, iid returns alongside large sample sizes may force out of date and therefore irrelevant information into the forecasts. Controlling for the empirical properties of the returns requires an augmented set of specifications. For instance, the AutoRegressive Moving Average (ARMA) Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) approach treats efficiently the returns time variation. However, the augmented parameter vector may induce estimation risk in the forecasts. In general, complex structures carry more model risk given that they are based on layers of models and distributional assumptions.

In this paper, we implement a variety of individual methods in order to evaluate their performance and create the combination methods pool of forecasts. In more detail, with respect to the parametric approach we implement a normal distribution unconditional fit, JP Morgan's univariate Riskmetrics method and members of the ARMA-GARCH family models.[4] Specifically, we focus on first order volatility processes based on the ARCH, GARCH and EGARCH specifications. Although the aforementioned specifications are members of the same family, each one of them has its own characteristics. For instance, GARCH specifications provide a well documented all around volatility approximation while ARCH specifications allow for less volatility persistence. On the other hand, EGARCH specifications incorporate the returns leverage effects and therefore account for possible skewness in the data. Finally, we employ the Normal, Student-t and Generalized Pareto (GP) distributions for the theoretical distribution of the error term .

With respect to the non-parametric approach, we implement three versions of the Historical Simulation (HS) method based on different sample sizes. Finally, from the hybrid approach we implement the Filtered Historical Simulation (FHS) and a Monte Carlo simulation based on the ARMA-GARCH(1,1)-t modeling specification. In total, we employ 14 individual forecasting methods, covering the literature's and financial practitioners' benchmark models. Appendix A provides a more detailed description of the implemented methods.

2.2 Combination Methods

The individual methods take into account certain empirical properties which in turn lend to each method specific characteristics and weaknesses. So the question is whether appropriate combination of these methods can eliminate inefficiencies and ultimately produce improved results compared to the individual ones. Timmermann (2006) suggests that combination estimators display superior performance because they combine the information embedded into each individual estimator. In addition, they are less sensitive to structural breaks and model mis-specification. Profits from combination forecasts also arise from diversification gains and correction for bias of individual forecasts.

[4] Riskmetrics and Historical Simulation are the financial practitioners' benchmark risk models.

So far the macroeconomic and financial literature suggests that the combination of individual forecasts produce superior results (see for example Stock and Watson, 2004; Huang and Lee, 2010; Rapach *et al.,* 2010). In the field of combination forecasts and especially volatility forecasts, the most recent studies (Becker and Clements, 2008; Patton and Sheppard, 2009; Andreou *et al.*, 2012) conclude that the combination methods outperform the individual ones. Specifically Becker and Clements (2008) investigate the performance of the S&P 500 Implied Volatility forecasts relative to model based forecasts and their combinations. Their results suggest that the combinations of model based forecasts are superior. In the same vein, Patton and Sheppard (2009) investigate the performance of Realized Volatility combination estimators relative to the individual ones and find that the simplest combination schemes produce superior results in most of the cases. Finally, Andreou *et al.* (2012) address the issue of model uncertainty in volatility by using a comprehensive model space and investigate whether a combination framework can improve volatility forecasts. More precisely, the authors consider the simple AutoRegressive models of Realized Volatility (AR-RV), the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV) (Corsi, 2009) and the Leverage HAR-RV (LHAR-RV)(Corsi and Reno, 2009) in addition to GARCH-type and nonparametric models of volatility. Their results suggest that forecast combinations based on an homogeneous robust loss function significantly outperform simple forecast combination methods, especially during the period of the recent financial crisis.

Risk management practices during the recent financial crisis is the focus of two recent papers closely related to our approach. McAleer *et al.* (2013) suggest using a combination of VaR forecasts to obtain a crisis-robust risk management strategy for a variety of international stock market indices. The authors develop VaR forecasts using combinations of the forecasts of individual VaR models, namely the r-th percentiles of the VaR forecasts of a set of univariate conditional volatility models. Their findings suggest that the median of the point VaR forecasts is a superior risk management strategy compared to strategies based on single and composite model alternatives. Relaxing the assumption of deterministic weights on individual forecasts, Halbleib and Pohlmeier (2012) propose a new methodology of computing VaRs based on the principle of optimal combination that accurately and robustly forecasts losses during periods of high risk. They consider two ways of computing optimal weights; first by minimizing the distance between the population and empirical moments derived from Basel II rules and other VaR evaluation techniques and second by employing simple quantile regressions on stand-alone VaR forecasts. Their combination forecasts exhibit a stable in-sample and out-of-sample performance across both calm and turbulent evaluation periods.

Our approach is more related to McAleer *et al.* (2013). In the same spirit, we aim at dealing with the adverse effects of the recent financial crisis on the potential mis-specification and inaccuracy of individual VaR methods by appropriately combining them. More precisely, we employ the mean, median and trimmed mean combination schemes of a large number of individual VaR forecasts. From an empirical point of view, the risk associated with the selection of a specific individual model (model risk) is eliminated, volatility of forecasts is reduced and structural instability is minimized. Moreover, our approach does not suffer from estimation error since weights are not estimated and is relevant to a wider audience due to its computational tractability.

Specifically, let us denote with $\widehat{VaR_{it}}, i = 1, 2, ..., n$ the n individual VaR methods. Then the mean, trimmed mean and median combining schemes for the VaR forecasts are the following:

$$\widehat{VaR}_{Mean,t}(q) = \frac{\sum_{i=1}^{n} \widehat{VaR_{it}}}{n} \tag{2}$$

$$\widehat{VaR}_{Trimmed1,t}(q) = \frac{\sum_{i'=1}^{n-1} \widehat{VaR_{i't}}}{n-2} \tag{3}$$

$$\widehat{VaR}_{Trimmed2,t}(q) = \frac{\sum_{i'=2}^{n-2} \widehat{VaR_{i't}}}{n-4} \tag{4}$$

$$\widehat{VaR}_{Median,t}(q) = Median(\{\widehat{VaR_{i't}}\}_{i'=1}^{n}) \tag{5}$$

where $\widehat{VaR_{i't}}$ denotes the sorted, in an ascending order, individual estimates.

3 Evaluation Framework

In order to evaluate individual forecasts and compare them with the combination methods, we use a two-step evaluation framework. First, we examine in detail the statistical accuracy of each method by applying a battery of backtesting criteria. Secondly, we evaluate the forecasting performance by utilizing the newly proposed methodology of Sener *et al.* (2012) in order to rank the VaR methods. These steps are considered independently while the final results are produced by ranking only the statistically accepted methods.

3.1 Statistical Evaluation

To statistically evaluate the employed VaR methods we use three different approaches plus an additional test. First we employ the approach proposed by Christoffersen (1998) which consists of three criteria, the unconditional coverage (LR_{uc}), the independency of the violations (LR_{ind}) and the conditional coverage criteria (LR_{cc}). To perform these tests a violations sequence is defined as follows:

$$I_t(q) = \begin{Bmatrix} 1, \ if \ r_t < -VaR_t(q) \\ 0, \ if \ r_t > -VaR_t(q) \end{Bmatrix} \tag{6}$$

The unconditional coverage criterion tests whether the empirical violations are consistent with the expected ones. Thus, this criterion tests the null hypothesis $H_{0,uc}$:$E[I_t(q)] = q$ against the $H_{1,uc}$:$E[I_t(q)] \neq q$. This hypothesis test is based on the following likelihood ratio test:

$$LR_{uc} = -2ln(\frac{(1-q)^{T_0} q^{T_1}}{(1-\frac{T_1}{T})^{T_0}(\frac{T_1}{T})^{T1}}) \sim \chi_1^2 \tag{7}$$

where T is the number of out-of-sample observations, T_0 the number of non violations and T_1 the number of violations. The independence criterion tests the interrelationship between the empirical violations. This test is performed using the following likelihood ratio:

$$LR_{ind} = -2ln((1-\frac{T_1}{T})^{T_0}(\frac{T_1}{T})^{T1}) + 2ln((1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}} \sim \chi_1^2 \quad (8)$$

where T_{ij} with $i, j = 0$ $(noviolation)$, $1(violation)$ is the number of observed events with the j event following the i event. The estimates of the probabilities of T_{ij} are marked as π_{01} and π_{11}. The conditional coverage criterion is essentially a synthetic criterion which tests the unconditional coverage and independence simultaneously. This test is performed using the following likelihood ratio:

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi_2^2$$

Another approach of evaluating the statistical performance of the VaR methods is the Engle and Manganelli (2004) Dynamic Quantile (DQ) approach. This method is based on a quantile regression model, by which the observed violations are associated with past violations and past information according to the following procedure:

$$Hit_t(q) = I_t(q) - q \quad (9)$$

$$Hit_t(q) = \left\{ \begin{matrix} 1-q, if\ r_t < -VaR_t(q) \\ -q, if\ r_t > -VaR_t(q) \end{matrix} \right\} \quad (10)$$

$$Hit_t(q) = \delta + \sum_{j=1}^{K} \beta_j Hit_{t-j}(q) + \sum_{j=1}^{K} \gamma_j \zeta_{t-j} + \varepsilon_t \quad (11)$$

where $Hit_t(q)$ denotes the modified violations sequence, δ is a constant term and ζ_{t-j} the variable that corresponds to any information that can be derived from the existing sample of observations. The null hypothesis of independence, DQ_{ind}, dictates that $\beta_j = \gamma_j = 0$, $\forall j = 1...K.$, while the null unconditional coverage hypothesis, DQ_{uc}, dictates that $\delta = 0$. Therefore, the null hypothesis of the conditional coverage criterion DQ_{cc} is defined as follows:

$$DQ_{cc} = (\frac{\Psi' Z' Z \Psi}{q(1-q)}) \sim \chi_{2K+1}^2 \quad (12)$$

where Ψ is the $2K + 1$ parameter vector and Z denotes the matrix of the regression variables. In this paper we set the maximum number of K lags equal to 3.

Another approach of evaluating the statistical performance of the VaR methods is the Christoffersen and Pelletier (2004) duration approach. While the Christoffersen and Engle-Manganelli approaches view the violation as the event, the duration approach takes into account the time interval between two violations. In other words, it evaluates the independence and conditional coverage hypothesis by testing the statistical properties of the sequence of time intervals between violations. The duration approach is based on the rationale of dependence causing violations clusters. In order to perform the tests, we define d_v as the time interval between the $v-1$ and the v violation. In the conditional coverage case, the evaluated method should produce exactly q violations equally time distanced. Consequently, the violation sequence will be characterized by a distribution with no memory. This entails that the d_v sequence will follow the Geometric distribution, i.e.

$$f(d_v, q) = q(1-q)^{d_v - 1}, d_v \in N \tag{13}$$

It is obvious that this evaluation method is based on determining a no memory distribution for the time intervals between the violations. In other words the probability of a violation at time t does not depend on the number of days that have passed since the previous violation. The only continuous distribution which is characterized by lack of memory is the Exponential distribution given by the following formula:

$$f(d_v, q) = qe^{-qd_v}$$

Christoffersen and Pelletier (2004) considered the Weibull distribution with parameters a, b for the alternative hypothesis of the non conditional coverage;[5]

$$w(d_v, a, b) = a^b b d_v^{b-1} e^{-ad_v^b} \tag{14}$$

The duration approach tests of independence and conditional coverage are converted in a Weibull parameter estimation procedure. Consequently the null hypothesis of the independence (Dur_{ind}) is not rejected if $b = 1$. In addition the conditional coverage (Dur_{cc}) hypothesis is not rejected if $b = 1$ and $a = q$.

The aforementioned unconditional coverage tests do not account for the severity of the violations. This entails significant risk especially in the case of extreme losses. To address this shortcoming, Colletaz *et al.* (2013) proposed a test which takes into account the severity of each violation. Specifically, a second violation sequence is defined as follows:

$$J_t = \begin{Bmatrix} 1, if\ r_t < -VaR_t(q') \\ 0, if\ r_t > -VaR_t(q') \end{Bmatrix},\ q' < q \tag{15}$$

where q' is a stricter significance level. The second violations or super exemptions sequence aims at measuring the number of initial violations that exceed the second threshold. Thus, if the VaR method produces an acceptable number of violations in conjunction with an increased number of super exemptions, the null hypothesis of the test will be rejected. To perform the test, three indicator functions are introduced:

$$g_{0,t} = 1 - g_{1,t} - g_{2,t} = 1 - I_t$$

$$g_{1,t} = I_t - J_t = \begin{Bmatrix} 1, if\ \ -VaR_t(q') < r_t < -VaR_t(q) \\ 0, if\ \ r_t < -VaR_t(q') \end{Bmatrix}$$

$$g_{2,t} = J_t = \begin{Bmatrix} 1, if\ r_t < -VaR_t(q') \\ 0, if\ r_t > -VaR_t(q') \end{Bmatrix}$$

[5] The Exponential distribution can be derived from the Weibull distribution for $b = 1$.

The above random variables follow the Bernoulli distribution with $1-q$, $q-q'$, q' parameters respectively. The null hypothesis of the test is defined as $H_{0,muc}$:$E[I_t(q)] = q$ and $E[J_t(q')] = q'$ and the test is performed via the following likelihood ratio:

$$LR_{muc} = -2ln((1-q)^{N_0}(q-q')^{N_1}\left(q'\right)^{N_2}) + 2ln((1-\frac{N_0}{T})^{N_0}(\frac{N_1}{T})^{N_1}(\frac{N_2}{T})^{N_2}) \sim \chi_2^2 \quad (16)$$

where $N_{i,t} = \sum_{t=1}^{T} g_{i,t}, i = 0, 1, 2.$

3.2 Performance Evaluation

The statistical evaluation tests sort the VaR methods according to their ability to produce the correct number of uncorrelated violations. However, they cannot provide information about the performance of the methods both in terms of underestimation or overestimation of the required capital. Another issue that arises is which method is the appropriate one in the case of multiple methods meeting the statistical criteria. To address these issues numerous loss functions and tests have been proposed.[6] In this paper we adopt the newly proposed approach of Sener *et al.* (2012) in order to rank the VaR methods. The main advantage of this approach is that it allows for a weighting between the underestimation and overestimation error. In order to rank the methods, penalization functions grade the underestimation error $(r_t - (-VaR) < 0)$ and the overestimation error $(r_t - (-VaR) > 0)$. These quantities represent, for a long position, the unexpected loss and the excessive allocated capital, respectively.

For the violation space, i.e. for all the violations, clusters of successive violations are constructed in order to compute the severity of the unexpected loss associated with each string of violations. For non-successive violations, the clusters are single elements. Let us assume that the $i-th$ cluster has z_i successive violations. The severity of unexpected losses for the $i-th$ cluster of violations is calculated as follows:

$$C_i = \prod_{b=1}^{z_i} (1 + (VaR_{b,i} - r_{b,i})) - 1, i = 1, 2, ..., \alpha \quad (17)$$

where $VaR_{b,i} - r_{b,i}$ denotes the error corresponding to each violation (for $b = 1, 2, ..., z$) of the $i-th$ cluster and α is the number of clusters. In addition to the severity of each cluster, the proposed methodology takes into consideration the correlation between clusters. Therefore, the penalization function for the unexpected loss is calculated as:

$$\begin{aligned} \Phi\left(VaR, r_t\right) &= \sum_{i=1}^{\alpha-1}\sum_{m=1}^{\alpha-i} C_i C_{i+m} \quad (18) \\ &= \sum_{i=1}^{\alpha-1}\sum_{m=1}^{\alpha-i} \frac{1}{d_{i,i+m}} \left(\prod_{b=1}^{z_i} (1 + (VaR_{b,i} - r_{b,i})) \prod_{b=1}^{z_{i+m}} (1 + (VaR_{b,i+m} - r_{b,i+m})) - 1 \right) \end{aligned}$$

[6] See for example, Angelidis and Benos (2008) who employ standard forecast evaluation methods in order to examine whether any differences between competing models are statistically significant.

where $d_{i,i+m}$ is the distance in days between the clusters. The penalization function of overestimation errors is derived in a similar way. However, given that the correlation between the excessive allocations does not affect directly the investor's position, it is not important for the computation of the excessive allocated capital penalization function. The penalization function for the overestimation of VaR is calculated :

$$\Psi\left(VaR, r_t\right) = \sum_{t=1}^{T} \left[I\left(r_t > -VaR | r_t < 0\right)\right]\left(\ VaR - r_t\right) \tag{19}$$

The penalization measure is calculated as:

$$PM\left(\Theta, r_t, VaR\right) = \frac{1}{T^*}\left[\left(1-\Theta\right)\Phi\left(VaR, r_t\right) + \Theta\Psi\left(VaR, r_t\right)\right] \tag{20}$$

where Θ denotes the weight on each penalization function and T^* denotes the number on non positive returns. For this paper Θ has been set equal to the VaR estimates significance level (q). The PM ratio for the j-th method is defined as:

$$PM_{j-ratio} = \frac{PM_j}{\sum_{j=1}^{n} PM_j} \tag{21}$$

The method with the lowest ratio is the best performing one. In addition to the ranking methodology, Sener *et al.* (2012) propose a supplementary test in order to test statistically the equality of performances between candidate methods. The main advantage of this test is the use of a ratio which eliminates the use of a benchmark method and reduces the computation complexity. For the supplementary test the PM ratio is set as the loss function. The loss series for the j-th method is defined as:

$$\left\{k_{j,t}\right\}_{t=1}^{T} = \left\{\frac{PM_{j,t}}{\sum_{j=1}^{n} PM_{j,t}}\right\}_{t=1}^{T} \tag{22}$$

If all the methods perform equally well, the aforementioned ratio should be equal to $1/n$. Hence, the null hypothesis for the j-th method is defined as:

$$H_0 : E\left[k_{j,t}\right] \leq \frac{1}{n},\ j = 1, ..., n$$

and the test statistic is defined as:

$$W_j = \sum_{t=1}^{T} I(k_{j,t} > \frac{1}{n}) \tag{23}$$

$$\widehat{W_j} = \frac{W_j - pT}{\sqrt{p(1-p)}} \sim N(0,1) \tag{24}$$

where p is the probability of $k_{j,t} > \frac{1}{n}$, which is set equal to 0.5. The above procedure reduces significantly the computational complexity. If a number of methods produce relative high ratios, a false non rejection may occur. To remedy this disadvantage, we evaluate the methods' performance in two steps. First we calculate the PM ratios and the corresponding tests for all the implemented methods. Second, we exclude the methods with the highest PM ratio and perform the evaluation process again. Finally, it is worth noting that the Ratio ranking and the supplementary test may point to different directions.

4 Empirical Findings

4.1 Dataset and Descriptive statistics

The bulk shipping sector constitutes the major form of sea transportation. This sector is particularly fragmented with the most notable categories being the dry bulk sector, which involves raw non liquid materials transportation and the tanker sector which is for liquids. Each freight depends on the type of the contract,[7] the size of the vessel, the type of the cargo and the route followed. This heterogeneity makes freight rate dynamics more complex than those of traditional asset classes (stocks, stock indices, etc.). A major provider of these rates is the Baltic Exchange, which publishes daily various spot freight rates calculated by members (panelists) of the exchange.

The dataset used in this study consists of eleven indices of the Baltic Exchange, describing both the aggregate state of the market and individual routes of the dry and wet cargo sectors. The aggregate indices we employ are the Baltic Dry Index (BDI), the Baltic Panamax Index (BPI), the Baltic Capesize Index (BCI), the Baltic Dirty Tanker Index (BDTI) and the Baltic Clean Tanker Index (BCTI). BDI is the general index of the dry bulk freight market and is expressed in index points. It is calculated as the non equal weighted sum of BCI, BPI and BHMI (Baltic Handymax Index) time charter averages.[8] These sub-indices are composite dry-cargo indices across international routes that correspond to vessels of size 30-49,999, 50-79,999 and over 80,000 deadweight tones (dwt henceforth) for the BHMI, BPI and BCI, respectively. With respect to the tanker shipping sector, we employ the Baltic Dirty Tanker Index (BDTI) and the Baltic Clean Tanker Index (BCTI), which are the two major sub-sector indices of this market. The distinction between 'clean' and 'dirty' comes from the type of cargo and specifically from whether the oil product carried can be classified as clean or dirty.[9]

The aforementioned indices are aggregate indices and as such we expect their behavior to be smoother compared to their constituents. However, as they are averages of individual route or time charter contract indices, they do capture the general dynamics of the fragmented freight market and can be thought of as portfolios of freight rate positions covering large fleets of vessels. On the other hand, individual route indices are more relevant for market risk exposures

[7]The type of contract entails specific characteristics for the chartering of a vessel. There are four types of contracts, voyage, bareboat, time charter and contract of affreightment (COA). The voyage type consists of the simplest contract where the vessel is chartered for a specific voyage and the rate is calculated by the weight of the goods. Bareboat type contracts lease only the equipment while Time Charter contracts lease the equipment and the corresponding services. Finally, COA consists of a series of voyage contracts.

[8]The Baltic Handymax index was replaced by the Supramax index in 2006.

[9]Clean products consist of lighter (sweet) distillates, such as gasoline and kerosene, which are usually shipped via vessels with coated tanks to ensure the cleanliness of the product. Dirty products involve lower distillates and residual oil which is usually shipped in conventional tankers.

of smaller companies since these companies employ vessels in one of these routes which also serve as the underlying assets of freight rate derivatives. To this end, we employ two popular averages of the time chartered indices included in the calculation of BPI and BCI; namely the 4 Time Charter Average Capesize (4TC AVG CAPE) and the 4 Time Charter Average Panamax (4TC AVG PANAM). They are averages of specific routes, expressed in US dollars and measure the cost to hire the vessel per day.[10] We complement our dataset with four popular Worldscale routes of the Baltic Dirty Index;[11] namely route TD3 (Middle East Gulf to Japan, for Very Large Crude Carriers (VLCC) vessel sizes of 250,000 dwt), the TD5 route (West Africa to US Atlantic Coast (USAC), for Suezmax vessel sizes of 130,000 dwt), route TD7 (North Sea to Continent, for Aframax vessel sizes of 80,000 dwt) and route TD9 (Caribbean to US Gulf, for Panamax vessel sizes of 70,000 dwt). The choice of these routes and time charter averages stems from their liquidity in the freight forward and option markets. For our analysis, we concentrate on the behavior and analysis of the aggregate sector indices. The analysis and findings for the individual routes are presented and discussed in Section 5.

Our dataset covers the period from 1/11/1999 to 13/03/2012 including 3091 daily observations. The dataset is obtained from Clarkson's Shipping Intelligence Network. Table 1 presents the descriptive statistics of the daily logarithmic returns of the indices at hand. All series display means statistically equal to zero as it is typical with daily returns of financial assets. Moreover, all return series are highly volatile and leptokurtic. The tanker sector displays higher kurtosis and is more negatively skewed compared to the dry bulk sector. These features indicate an increased probability for extreme events and given the negative asymmetry (leverage effect) in the empirical distribution of the aggregate indices (with the exception of BCI) the odds are in favour of negative outcomes. In contrast to aggregate indices, individual routes exhibit positive skewness with the exception of 4TC AVG PANAM. As expected, the null hypothesis of normality is rejected as indicated by the Jarque-Bera test.

Contrary to more traditional asset classes (stocks, stock indices, etc.) freight rate returns display significant autocorrelation according to the Ljung Box statistic. This departure from the martingale difference hypothesis can be attributed to the specific characteristics of the freight rate market. For instance, Stopford (2009) argues that freight market equilibrium is conditional on the magnitude of the demand and supply fluctuations, while financial market equilibrium can be characterized as instantaneous. Furthermore, the freight rate market is strongly correlated to exogenous variables such as crude oil or mineral demand, economy growth etc. pointing towards augmented information sets. Finally, the return series exhibit significant heteroscedasticity as suggested by the Ljung Box statistic for the squared return series. These characteristics of the freight return series call for a VaR methodology that can adequately capture tail risk. The following sections are focused solely on the returns information set and leave the inclusion of

[10] Admittedly, these indices do not represent individual routes. They represent averages of individual ones. However, we loosely refer to them as 'individual' routes since they are more disaggregate indices.

For a description of the specific time chartered routes, see Table 1 in Angelidis and Skiadopoulos(2008).

[11] Worldscale is a freight rate measurement. It is used in the tanker sector and is calculated every year by the World Scale Association. Each rate is quoted as a percentage of 100 Worldscale.

exogenous variables to future research.[12]

[TABLE 1 AROUND HERE]

4.2 Statistical Evaluation

Before presenting our findings, it is useful to briefly describe our forecast construction methodology. Given a total sample of K observations, we must determine the way to split the sample into the estimation part (say R observations) and the out-of-sample part (say $T := K - R$ observations). Obviously, there is a trade-off, since a large R improves the quality of the estimated parameters of the model but, at the same time, leaves few observations for the out-of-sample forecast exercise making the evaluation of the predictive ability of the model difficult. In our analysis, we keep about 1/3 of the available sample (1029 observations) for out-of-sample forecasting. This out of sample specification gives us a sufficient number of forecasts to evaluate the estimated models, while keeping enough observations to obtain reliable in sample parameter estimates. We employ a rolling forecasting scheme, i.e. the size of the estimation sample remains fixed (equal to $R = 2061$ observations) and produce VaR estimates for the 5% and 1% significance levels. These forecasts are used to conduct both the statistical and the performance evaluation.

Table 2 presents the detailed BDI backtesting results by providing the p-values associated with the statistical evaluation of our methods (see Section 3.1). Table 3 summarizes the results for the remaining aggregate indices in number of rejections for every test category.[13] More in detail, for the BDI index and the 5% estimates (Table 2, Panel A) our results suggest that the hybrid and parametric GARCH-type methods (AR-GARCH(1,1)-N, AR-GARCH(1,1)-t, AR-EGARCH(1,1)-N, AR-EGARCH(1,1)-t) pass all the conditional coverage criteria, while the Historical SimulationAll and Variance Covariance methods fail on the basis of all tests. Interestingly, Historical Simulation-500 and Historical Simulation-250 are not rejected by any unconditional coverage test specification. The high volatility periods of the index combined with the small in-sample specification inflate (overestimate) the VaR forecasts, rendering a small number of violations. However, these violations cluster and therefore point towards dependency of VaR forecasts on past information. Finally, the Combination Median method succeeds in meeting all the backtesting criteria, while the remaining combination methods fail to respond adequately to the violation clustering. The failure of the mean and trimmed mean combination methods are in a sense expected, since almost half of the individual methods exhibit non tolerance to violation clustering.

[TABLE 2 AROUND HERE]

Our findings with respect to BPI 5% VaR case (Table 3, Panel A) point to the superiority of the AR-ARCH(1)-N model which succeeds in all the backtesting criteria. However, most of

[12] We thank an anonymous referee for pointing this out. For each series at hand, we employ the appropriate AR specification based on the autocorrelation function. Specifically, we implement the following AR specifications for the mean equation: BCI: AR(2), BCTI: AR(3), BDI: AR(2), BDTI: AR(2), BPI: AR(2), 4 TC Av Ca: AR(2), 4 TC Av Pa: AR(2), TD3: AR(1), TD5: AR(1), TD7: AR(1), TD9: AR(1).

[13] The detailed tables are given in Appendix B, which is available from the authors upon request.

the parametric GARCH and hybrid methods meet the conditional coverage criteria. For the combination methods we find similarities with the individual methods' results. Specifically, only the Combination Median produces statistically accurate forecasts while the remaining combination methods succeed in specific unconditional coverage tests. Turning to the BCI Index, our findings lead to similar results pointing to the superiority of parametric (GARCH type) and hybrid methods. In addition to GARCH volatility modelling, the Combination Median method meets all the criteria while the trimmed mean methods are successful in at least two conditional coverage criteria (Standard approach and Duration approach).

Our findings for the tanker sector paint a different picture. Specifically, with respect to the BDTI 5% VaR case, only the Filtered Historical Simulation method passes all the conditional coverage criteria. All the remaining individual methods, with the exception of EVT, exhibit fewer violations which eventually lead to the rejection of the Conditional Coverage Hypothesis. Similarly, the combination methods overestimate VaR and lead to a significantly reduced number of violations. However, the combination methods produce very good results in terms of independence, indicating their suitability to correctly capture the violation clustering. Similar findings pertain to the case of BCTI with the addition of three more statistically accepted individual methods. More in detail, for the Clean tanker sub-sector the Filtered Historical Simulation method passes all the backtesting criteria while the Monte Carlo Simulation passes all the conditional coverage criteria. Finally, the AR-GARCH(1,1)-t and the AR-EGARCH(1,1)-t models pass two conditional coverage criteria. Despite the improved statistical performance of some individual methods, the performance of the combination methods does not show any signs of improvement. As in the BDTI case, the overestimation of market risk leads to a decreased number of violations.

[TABLE 3 AROUND HERE]

Moving to the 1% significance level, we find that the majority of individual methods are statistically rejected while the combination methods display superior performance. For the BDI index (Table 2, Panel B) we find that only the AR-GARCH(1,1)-t and AR-EGARCH(1,1)-t models pass all the backtesting criteria while the Filtered Historical Simulation method fails only on the basis of the LR_{muc} test. For the combination methods we find that the Combination Mean and Combination Trim1 methods pass all the criteria while the Combination Median and Combination Trim2 methods are rejected by one conditional coverage test. As far as the BPI index is concerned (Table 3, Panel B), mainly the parametric and hybrid methods produce statistically accurate forecasts with the AR-GARCH(1,1)-t, AR-EGARCH(1,1)-t and hybrid methods passing all the criteria. The Combination Median method emerges as the best among the combination methods followed by the Combination Mean and Combination Trim2. Our findings regarding the BCI index are similar to the BDI index, where the AR-GARCH(1,1)-t, AR-EGARCH(1,1)-t and hybrid methods pass all the backtesting criteria. More importantly, all the combination methods succeed in all the backtesting criteria.

For the tanker sector, our findings are quite different compared to the 5% significance level. For BDTI, eleven of the individual methods pass at least two conditional coverage criteria, while six pass all the backtesting criteria. In addition, the Variance Covariance and Historical Simulation methods are not rejected by every specification. In addition, HS 250 does not fail any test. Similarly to the BDI 5% results, the BDTI increased volatility and excessive kurtosis in conjunction with the small out of sample period leads to overestimation of risk. However, the number of violations at the 1% coverage level is small, making violation clustering more difficult. This is supported by the rejections of the HS methods for the BDTI 5% coverage level. More importantly, the combination methods display an improved performance. Specifically, the CombinationTrim2, Combination Mean and Combination Median methods fail only on the basis of the LR_{muc} test. With respect to the Combination Trim1 method, we should clarify that its failure at the LR_{uc} and DUR_{uc} test is attributed to the overestimation of risk. In addition, the successes at all the DQ_{cc} and independence tests, the borderline failure at the LR_{cc} and finally the second highest LR_{muc} p-value suggest that Combination Trim1 can be considered as a statistically adequate method. The results for the BCTI Index are quite similar but with fewer successes for the individual methods. All the parametric and hybrid methods pass at least two conditional coverage criteria while the Historical Simulation-250, Historical Simulation-500 and the Variance Covariance methods fail. With respect to the combination methods, we find an improved performance compared to BDTI since they all pass at least two conditional coverage criteria and three of them pass all the conditional coverage criteria.

4.3 Performance Evaluation

While the statistical evaluation framework provides a way of examining the statistical accuracy of the VaR methods, it does not offer any insight on the economic performance of the alternative methods with the exception of the Colletaz *et al.* (2013) test. In other words, the statistical backtesting framework does not offer a measure of overestimation or underestimation of market risk, which in turn is crucial in selecting the most reliable method. In this paper, we employ the performance evaluation approach proposed by Sener *et al.* (2012), described in detail in Section 3.2. The evaluation proceeds in two steps. First, we derive the loss ratio and the supplementary test for all the methods. Our findings are reported in Table 4 (Panels A and B for the 5% and 1% VaR, respectively). At a second step, we repeat the analysis by excluding the worst performing methods, i.e. those with significantly high ratios. In this way, we make sure that we get more accurate results which are not affected by extreme ratios.

[TABLE 4 AROUND HERE]

For the BDI index and the 5% VaR level, the Penalization Measure (PM) (Equation 20) ranges from 0.135 for the AR-EGARCH(1,1)-N model to 0.592 for the HistoricalAll simulation method. The corresponding ratios (Equation 21) range from 3.255 to 14.265. Overall, our findings suggest that the non-parametric Variance Covariance methods and the AR-ARCH(1)-t methods are outperformed by the parametric, hybrid and combination methods. With respect to the combination methods, we have to note that they belong to the pool of the equally performing methods with the Combination Median performing best. Excluding the methods with the relative worst PM ratios (Table 5, Panel A), we find that the AR-ARCH methods, the Riskmetrics and the EVT are outperformed by the remaining methods while the initial ranking is maintained. At the 1% VaR level (Table 5, Panel B) our findings are similar with the Variance Covariance, the AR-ARCH(1)-N and most of the non-parametric methods being outperformed by the remaining methods. However, at the 1% significance level the ranking of the accepted methods differ in comparison with the 5% level. Specifically the three highest performing methods are the mean combination methods while the median ranks fifth. Excluding the worst ranked methods (Table 5 Panel B), we find that the Riskmetrics, the Historical Simulation-500 and three of the parametric methods are statistically out performed while the equal performing methods maintain their ranking.

[TABLE 5 AROUND HERE]

For the BPI index, our 5% VaR level results are generally similar to the BDI index. The majority of the parametric and hybrid methods outperform the non-parametric ones, the AR-ARCH(1)-t and Variance Covariance methods. On the other hand, the combination methods cannot be rejected with the median method ranking third. Excluding the worst performing methods, we find that the GARCH-type methods, the Filtered Historical Simulation, the Monte Carlo Simulation, the Median and Trim2 Combination emerge as superior methods. Moving to the 1% VaR level (Table 4, Panel B) we find that the Historical SimulationAll, the Historical Simulation-500 and the Variance Covariance methods are the worst performing ones, closely followed by the Historical Simulation-250, the Riskmetrics, the AR-ARCH(1)-N and the Combination Mean method. Excluding the worst performing methods (Table 8, Panel B) we find that the pool of rejected methods is joined by the Historical Simulation-250, Riskmetrics, AR-ARCH-t, Mean and Trim2 combination methods.

Turning to the BCI index (Table 4, Panel A), our findings suggest that at the 5% significance level two of the nonparametric methods, the Variance Covariance plus the AR-ARCH methods are outperformed by the remaining ones. In addition, we find that the Combination Median method ranks first and the alternative combination methods rank from the fourth to the sixth place. The re-evaluation of the methods (Table 5, Panel A) suggests that no specific group of methods is superior while the rejected methods consist of the non-parametric Historical Simulation-500, the Riskmetrics, the parametric EVT and the AR-ARCH(1)-N methods. At the 1% VaR level (Table 4, Panel B) our findings are similar. The combination methods produce mixed signals since the Combination Trim1 method ranks first while the remaining combination methods rank from the seventh to the ninth position. The second step evaluation (Table 5, Panel B) does not yield any differences in the ranking with the combination methods being statistically equal to the best performing methods.

For the BDTI index at the 5% significance level, we find that the non-parametric, the Variance Covariance, the EVT and AR-EGARCH(1,1)-N are outperformed by the remaining methods, while the combination methods rank within the first five places with the Combination Mean ranking first. Our second step evaluation findings suggest that the combination methods perform better than any other group of methods. Four of the parametric methods and one of the hybrid ones are outperformed by the remaining methods which in general keep their first step evaluation ranking. Our 1% significance level results show that the rejected methods are the same as previously with the exception of the EVT and Historical Simulation-250. The second step evaluation (Table 5, Panel B) does not yield any substantial differences with the exception of the Historical Simulation-250 forecasts equality of performance rejection. This finding suggests that the model's statistical accuracy is based on the overestimation of risk. Although statistically accurate, the model cannot compete against parametric or hybrid methods. Finally, the combination methods do not perform as well as at the 5% significance level despite the third place of the Trim1 method and the fact that only the median method is rejected.

For the BCTI index and the 5% VaR level (Table 4, Panel A) six of the individual methods are rejected including the non-parametric methods, the Variance Covariance and representatives from the other groups of methods with the exception of the combination methods. Specifically, the combination methods occupy the four top positions with the Combination Median method ranking first. Excluding the worst performing methods (Table 5 Panel A) we find that the combination methods and the parametric with normal realizations outperform the remaining methods. With respect to the 1% significance level (Table 4, Panel B) we show that except for the non-parametric methods and the Variance Covariance methods, the remaining methods perform equally well. Moreover, the first two positions are occupied by the Trimmed Combination methods while the Combination Median and Combination Mean rank sixth and seventh respectively. At the second step evaluation (Table 5, Panel B) we find that six methods are rejected including the parametric with normal errors, the EVT, the Riskmetrics and finally the Filtered Historical Simulation.

4.4 Joint Evaluation

The findings presented so far are indicative of the statistical and performance evaluation independently. In order to evaluate the overall ability of the methods to measure market risk accurately and efficiently, we have to combine the results of the statistical and performance evaluation. To this end, we rank the statistically accepted methods, presented in Section 4.2 by employing the performance evaluation ranking results presented in Section 4.3. As already mentioned, we consider as a statistically accepted method any method that has passed at least two conditional coverage criteria. Table 6 presents the ranking of the statistically accepted methods for the two significance levels.

For the 5% significance level (Table 6, Panel A) we could conjecture that due to the time-varying nature of volatility, the methods with a GARCH-EGARCH-Normal volatility mapping produce superior results compared with the non-parametric ones. Regarding the combination methods, the Combination Median performs equally well in every case while it ranks first in the BCI case. Moving to the Tanker sector our findings differ mainly due to the nature of the returns of the corresponding indices. Due to the overestimation of risk, the final evaluation of the methods for the BDTI and BCTI indices include the methods that overestimate market risk. The respective findings show that combination methods outperform the individual ones.

Overall the results in the dry and tanker sector suggest that the Combination Median can be considered as a global method of measuring market risk or at least as an accurate proxy for the expected daily 5% VaR of the freight rates.

[TABLE 6 AROUND HERE]

Our findings with respect to the 1% significance level (Table 6, Panel B) paint a different picture. For both sectors the combination methods rank amongst the top positions[14] alongside specific individual methods. Contrary to the 5% significance level, the parametric methods with GARCH-EGARCH-Student volatility mapping perform better than the non-parametric, other parametric methods and the hybrid methods. However, the very good performance of these individual methods cannot mask the superior performance of the combination methods and especially the Combination Trim1 method, which ranks first in three indices and third in one. At this point, we should mention that the failure in BPI is due to overestimation of risk which in turn leads to the failure of the unconditional coverage hypothesis. However the independence criteria point to the adequacy of the method, a fact that is supported by both the PM ratio and the performance equality test.[15]

To conclude, we find that the combination methods are a strong alternative to the large set of individual methods. At the 5% VaR, the Combination Median method is a globally accepted one. This method performs equally well as the highly-ranked individual methods and thus can significantly reduce the cost of freight rate risk measurement. For the 1% VaR, we find that the combination methods are superior to the individual methods with Combination Trim1 achieving the highest performing ranking in almost every case.

5 Robustness Checks: Individual routes

In order to examine the robustness of our main findings, we repeat our analysis for the six individual routes already described in section 4.1. Quite importantly, individual routes are practically relevant from the shipping industry perspective since ships can be fixed on a floating rate charter based on the daily value of the Baltic time charter averages, for example. Table 7 summarizes the backtesting results in number of rejections per test category, while Table 8 presents the ranking of the statistically accepted methods for the two significance levels.[16]

Starting with the time chartered averages, we note that their heteroscedastic and heavy tailed profile leads to failure of the majority of the individual methods (Table 7, Panel A). For the 4 TC AVG CAPE, only the AR-GARCH, AR-EGARCH and Monte Carlo Simulations methods prove adequate. Despite the failure of the majority of the individual methods, the Combination Median method produces excellent results succeeding in every backtesting criteria.

[14]The acceptance of the Combination Trim1 method as an accurate method is due to the borderline backtesting results (see section 4.2) in conjunction with the superior results in the evaluation ranking.

[15]In contrast with the Combination Trim1 in the BDTI case, for the BPI the combined results of the statistical and performance evaluation are not sufficient to surpass the "at least two Conditional Coverage criteria pass" rule and therefore consider the method as an adequate one.

[16]The detailed tables are given in Appendix B, which is available from the authors upon request.

Similar results are reported for the 4 TC AVG PANAM 5% statistical evaluation where the AR-ARCH(1)-N, AR-GARCH, Filtered Historical and Monte Carlo Simulation methods pass all the backtesting criteria. However, in this case none of the combination methods produce statistically accurate results. Turning to the individual tanker routes, our findings suggest that the majority of the parametric and hybrid methods are statistically accurate for the TD5, TD7 and TD9 routes. Quite importantly, the proposed combination methods produce excellent results. In contrast to the aforementioned three routes, only the student-t parameterized methods prove adequate for TD3, as the remaining methods overestimate risk.

Moving to the 1% statistical evaluation (Table 7, Panel B), the 4 TC AVG CAPE results suggest that there is a significant improvement in the statistical performance of the combination methods. Combination Mean, Combination Trim 1 and Combination Trim 2 produce statistically accurate results, while only the AR-GARCH(1,1)-t and the Monte Carlo Simulation qualify as statistically adequate individual methods. For the 4 TC AVG PANAM, the AR-GARCH-EGARCH and Hybrid methods, produce accurate results. In addition, three of the combination methods, namely Combination Mean, Combination Median and Combination Trim 2, produce accurate results with Combination Median passing all the backtesting Criteria. A similar improvement is also evident for the individual tanker routes, where the majority of the individual methods and all the combination methods produce accurate results, even for the TD3 case.

[TABLE 7 AROUND HERE]

The first step of the performance evaluation for both sectors produces results similar to the statistical evaluation ones. Specifically, at the 5% significance level, the 4 TC AVG CAPE results suggest that the parametric (AR-GARCH, AR-EGARCH), hybrid and combination methods perform equally well to the best performing method. Similarly, the 4 TC AVG PANAM results are almost identical with the addition of EVT, Riskmetrics and AR-ARCH(1)-N. For the Tanker Sector the student parameterization seems to underperform while in some cases (TD3, TD5) the non-parametric Historical Simulation-250 and Historical Simulation-500 perform well. In addition, the Riskmetrics is performing equally in every index while the Variance Covariance is not rejected only for the TD5 index. At the 1% significance level, our results are similar with the combination methodology producing non-rejection results in every case examined. The hybrid and parametric GARCH-EGARCH-Student methods produce a single rejection in the TD3 index while the Variance Covariance is rejected in every case. Finally, it is worth mentioning that there is at least one non parametric method performing equally well in every index.

The second step evaluation results are consistent with the first step results. At the 5% significance level, the Dry Bulk results are almost identical for the parametric AR-GARCH, AR-EGARCH and Hybrid methods. However, in the Panamax Sector there is a rejection of performance equality for the Combination Mean, Combination Trim 1 and Combination Trim 2 methods. At the tanker sector there are rejections of equal performance for the parametric methods with student-t parametrization, while the combination methodology performs equally in almost every case. An exception has to be made for the TD5 index where the Combinations Trim methods fail to perform equally well. At the 1% level evaluation, only the Filtered Historical Simulation presents a non rejection in every case while the Montecarlo and the AR-GARCH(1,1)-t methods produce a rejection in TD3 index. The remaining individual methods do not unveil a specific profile as there are random rejections and non-rejections. Finally, the performance of the proposed combination methodology is encouraging with at least one non-rejection in every case and four non rejection in three routes (4 TC AVG CAPE, TD3, TD5).

Combining the results of the statistical and the performance evaluation reveals that the combination methods constitute a strong alternative. Specifically, at the 5% significance level (Table 8, Panel A), there is not a single method globally accepted. Alongside the Combination Median, which is equally performing with the top ranked method in four indices, the AR-GARCH(1,1)-N produces similar results. In addition, our combination methods perform better in the tanker sector where for the TD5, TD7 and TD9 indices there are at least two methods in the best performers' group. Moving to the 1% significance level (Table 8, Panel B), our results are in favor of the proposed combination methodology. Specifically, in every case there is a statistically accurate best performing combination method with Combination Trim1 alongside AR-GARCH(1,1)-N presenting a global profile.

To conclude, measuring the market risk of the individual routes is quite challenging due to the specific characteristics attached to each one of these routes. Our findings imply that although there is not a definite superiority of the combination methodology, it represents a reliable alternative that reduces significantly model uncertainty and parameter instability contained in the individual methods. This is even more pronounced in the 1% VaR calculation where the combination methodology provides, in every case, at least one statistically accepted method performing equally well to the best performing one. These results confirm the superiority of the proposed combination methodology and the robustness of our approach. This finding reinforces its potential suitability for other markets such as the freight derivatives market and various financial markets.

6 Conclusions

Freight rate risk is one of the most important risk factors of the shipping industry. In addition, the profile of the distributions of returns complicates the measurement of freight risk leading to poor performances by the majority of the VaR methods. In this paper we provided a thorough insight of freight risk via a VaR methodology. Specifically, we considered the performance of the standard VaR methods amid the recent adverse economic circumstances while most importantly we proposed a Combination approach aiming at superior VaR results. In order to evaluate the employed models/ methodologies, we implemented a two step evaluation methodology. For the backtesting of the implemented methods, three distinct approaches plus a newly proposed test were applied in order to achieve an in depth statistical evaluation. Specifically, besides the standard approach proposed by Christoffersen (1998) we implemented the Dynamic Quantile approach proposed by Engle and Manganelli (2004), the Duration approach proposed by Christoffersen and Pelletier (2004) and finally the super exemption test proposed by Colletaz *et al.* (2013).

For the evaluation of the forecasting performance, we implemented a newly proposed methodology introduced by Sener *et al.* (2012). The main advantages of the this methodology, in contrast with the standard loss functions, are two. First it allows weights to each type of error, making it possible to distinguish between underestimation and overestimation errors. Second, it significantly reduces the computation complexity associated with the testing of performance equality.

Our findings suggest that on the basis of individual methods only the parametric and hybrid methods produced acceptable results and seem to adapt better to the volatile nature of the freight market. More importantly, we found that the combination methods produce better results than the individual methods posing a strong alternative to the large number of individual methods. In addition, we found combination methods that present a global profile applicable throughout the entire freight market. Therefore, the combination methods we propose can provide accurate results while simultaneously reducing the cost of model evaluation and selection.

Further research should investigate the possibility of expanding the model space to incorporate models based on realized volatility estimators, such as the AR-RV, HAR-RV and LHAR-RV models (Andreou *et al.*, 2012). In this respect intra-day data should be employed, which unfortunately are currently unavailable due to the nature of the spot freight rate market. In addition, such an approach can be pursued for the evaluation of risk in futures freight contracts for which it might be probable that IMAREX provides intra-day data in the near future. Finally, current revisions to the Basel III framework suggest moving from VaR to Expected Shortfall. In this respect, it would be worth investigating whether our framework can be of sufficient value in this new era for risk management practices.

Appendix A: Individual VaR Forecasting Methods

In our analysis, we utilize a variety of individual forecasting methods which can be divided into three categories; parametric, non-parametric and hybrid methods. The following sections provide a detailed description of the implemented specifications.

A.1. Parametric methods

The parametric approach requires the fit of a known theoretical distribution as an alternative to the real DGP. The underlying idea is to provide an approximation that would treat efficiently the returns empirical properties and at the same time would lay a structure for the risk measures calculations.

Unconditional Distribution Fit

The simplest parametric approach consists of a theoretical distribution fit directly to the raw empirical returns. It assumes iid returns and utilizes the sample estimates to infer the parameters in question. In more detail, let $r_1, ..., r_N$ be a random sample of returns. Then $r_t \sim F(\theta)$ $\forall t = 1, ..., N$ where $F(\theta)$ is the CDF of a theoretical continuous distribution with parameters vector θ. The Variance Covariance method assumes zero mean Normal iid returns and uses the sample standard deviation as a proxy for the volatility. The VaR estimates are produced as follows:

$$\widehat{VaR_t}(q) = F_q^{-1}(z, \theta)\widehat{\sigma} \tag{25}$$

where $F_q^{-1}(z, \theta)$ denotes the q quantile of the normal distribution and $\widehat{\sigma}$ is the in sample standard deviation.

Conditional Distribution Fit

A more elaborate parametric approach is to consider the time variation of returns and adjust the fitted distribution accordingly. This requires to model the distribution's parameters evolution in an attempt to control for time variation. Based on a mean variance modeling process, we can estimate VaR as follows:

$$\widehat{VaR_t}(q) = \widehat{\mu_t} + F_q^{-1}(z, \theta)\widehat{\sigma_t} \tag{26}$$

where $\widehat{\mu_t}$ is the mean return estimate, F_q^{-1} is the quantile of the standardized theoretical distribution given the estimated parameter θ, and $\widehat{\sigma_t}$ is the estimate of the conditional standard deviation. If the series of returns present autocorrelation, while stationary, the AutoRegressive Moving Average (ARMA) model approach is suitable for the mean estimation. The ARMA(P, Q) process is given by the following equation:

$$\mu_t = \varphi_0 + \sum_{i=1}^{P} \varphi_i r_{t-i} + \sum_{j=1}^{Q} \psi_j \varepsilon_{t-j} \tag{27}$$

where $\sum_{i=1}^{P} \varphi_i r_{t-i}$ denotes the dependence with lagged observations of returns and $\sum_{j=1}^{Q} \psi_j \varepsilon_{t-j}$ denotes the dependence of returns with lagged errors. To model volatility, we employ models that belong to the Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) family. Specifically, we employ the AutoRegressive Conditional Heteroscedastic (ARCH) model proposed by Engle (1982) which can be considered a special case of the GARCH models proposed by Bollerslev (1986). A GARCH(P, Q) model is given by the following equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{P} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{Q} b_j \sigma_{t-j}^2 \tag{28}$$

$$\alpha_0 > 0, \alpha_i \geq 0, b_j \geq 0 \; and \; \sum_{i=1}^{max(P,Q)} (\alpha_i + b_i) < 1$$

while ARCH models can be derived from equation (28) for $Q = 0$. While the above models can sufficiently capture the volatility of returns, they cannot discriminate between negative or positive prior returns. In order to capture the potential leverage effect, we employ the Exponential GARCH (EGARCH) models proposed by Nelson (1991). The EGARCH(P, Q) model is described by the following equation:

$$log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{P} (\alpha_i |\frac{\varepsilon_{t-i}}{\sigma_{t-i}}| + \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}) + \sum_{j=1}^{Q} b_j log(\sigma_{t-j}^2) \tag{29}$$

$$\alpha_0 > 0, \alpha_i \geq 0, b_j \geq 0 \text{ and } \sum_{i=1}^{max(P,Q)} (\alpha_i + b_i) < 1$$

In order to estimate the parameters of models (27)-(29), we employ both the Normal and Student-t distributions as the theoretical return distribution.

Finally, we follow the McNeil and Frey (2000) Extreme Value Theory (EVT) approach in order to simulate the tail of the empirical returns distribution and avoid the implementation of symmetrical distributions that may not account for the unique properties of the empirical distribution of returns.[17] The authors propose a two step approach that utilizes an ARMA-GARCH process in order to filter the historical returns and create iid error terms. Then a Generalized Pareto (GP) distribution is fitted on the error terms. With respect to the GP fit, we use the Hill estimator in order to attain the parameter of the GP distribution while setting as a threshold the 5% quantile of the sample of returns. The mean return is calculated as described in equation (27), while the standard deviation is estimated by a GARCH(1,1)-t model. The EVT-VaR is estimated as follows;

$$\widehat{VaR_t}(q) = \widehat{\mu_t} + \widehat{\sigma_t} u \left(\frac{q}{T_u/T} \right)^{(-\xi)} \tag{30}$$

$$\xi = \frac{1}{T_u} \sum_{i=1}^{T_u} \ln(\frac{y_i}{u}), y_i = r_i - u \tag{31}$$

where $\left(\frac{q}{T_u/T} \right)^{(-\xi)}$ is the q quantile of the GP distribution, u denotes the threshold and finally T and T_u are the number of sample observations included in the evaluation sample and the

[17]For a more detailed discussion on EVT theory see Rocco (2014).

calculation of the Hill estimator (Equation 31) respectively.

In conjunction to the aforementioned methods, we also implement the JP Morgan's univariate Riskmetrics method. Riskmetrics assumes normality of returns but rejects the homoscedasticity property. In addition, returns depend only on the past volatility as captured by an Integrated GARCH (IGARCH) model with fixed parameters. The VaR estimates are produced by the following specification:

$$r_t = \varepsilon_t, \varepsilon_t = z_t \sigma_t \tag{32}$$

$$\sigma_t^2 = 0.04\varepsilon_{t-1}^2 + 0.96\sigma_{t-1}^2 \tag{33}$$

$$\widehat{VaR_t}(q) = F_q^{-1}(z, \theta)\widehat{\sigma_t} \tag{34}$$

where $F_q^{-1}(z, \theta)$ denotes the q quantile of the normal distribution and $\widehat{\sigma_t}$ is the estimated conditional standard deviation.

A.2. Non Parametric Methods - Historical Simulation

In contrast to the parametric methods, the non-parametric ones calculate VaR using exclusively the empirical distribution of returns. The idea underlying the non-parametric processes is that of the repetition of past losses. The simplest and more popular member of the non-parametric methods is the Historical Simulation (HS) method. HS utilizes the iid returns assumption to calculate VaR as:

$$\widehat{VaR_t}(q) = F_q^{-1}(\{r_i\}_{i=1}^{t-1}) \tag{35}$$

where F_q^{-1} denotes the q quantile of the sample of returns.[18] By definition, HS incorporates the properties of the empirical distribution of returns and does not induce estimation risk. However, the iid assumption induces specification risk as it creates a sample average of all the events. In our study, we employ three versions of Historical Simulation estimates based on data samples of 250, 500 and T observations corresponding to a window of one year, two years and the whole sample. Given that the relevance of the information set is connected to the market conditions at the time of the forecast, restricting the information set may improve the performance of the HS VaR forecasts.

A.3. Hybrid Methods

In an attempt to combine both approaches and thus alleviate their disadvantages, Barone-Adesi *et al.* (1999) propose the Filtered Historical Simulation (FHS). FHS employs the mean and standard deviation procedure of the parametric methods (Equations (27)-(29)) in conjunction with the properties of the empirical distribution of returns. Essentially this method can be viewed as a parametric method without the theoretical distribution hypothesis. In this paper,

[18] HS is considered as a benchmark method for quantifying risk. Berkowitz and O'Brien (2002) and O'Brien and J.Szerszen (2014) report HS amongst the main risk measuring approaches of the largest United States financial institutions.

we model returns via an ARMA-GARCH(1,1)-t model and the VaR is calculated as follows:

$$\widehat{VaR_t}(q) = \widehat{\mu}_t + F_q^{-1}(\{z_i\}_{i=1}^{t-1})\widehat{\sigma}_t \tag{36}$$

where $F_q^{-1}(\{z_i\}_{i=1}^{t-1})$ denotes the q quantile of the standardized residuals. Similarly, we utilize a Monte Carlo approach where VaR is estimated as a quantile of a simulated empirical distribution of returns. Specifically an ARMA-GARCH(1,1)$-t$ model is employed in order to produce 10000 daily pseudo returns and VaR is produced as the q^{th} quantile of these simulated returns.

References

[1]Abouarghoub, W., Mariscal, I. B.-F., & Howells, P. (2014). A two-state Markov-switching distinctive conditional variance application for tanker freight returns. International Journal of Financial Engineering and Risk Management, 1, 239-263.

[2]Alizadeh, A., H. (2013).Trading volume and volatility in the shipping forward freight market. Transportation Research Part E: Logistics and Transportation Review, 49, 250-265.

[3]Alizadeh, A., H. , Huang, C.-H., & Dellen, S. (2014) A regime switching approach for hedging tanker shipping rates. Energy Economics, 49 (10), 44-59.

[4]Alizadeh, A., H., & Nomikos, N. (2009). Shipping derivatives and risk management. Palgrave Mcmillan, New York.

[5]Alizadeh, A., H., & Nomikos, N. (2011). Dynamics of the Term Structure and Volatility of Shipping Freight Rates. Journal of Transport Economics and Policy, 45 (1), 105-128.

[6]Andreou, E., Kourouyiannis, C., & Kourtellos, A. (2012). Volatility forecasts combinations using asymmetric loss functions. University of Cyprus, Working Paper in Economics, WP17-12.

[7]Angelidis, T., & Benos, A. (2008). Value-at-Risk for Greek Stocks. Multinational Finance Journal, 12 (1/2) 67–104.

[8]Angelidis, T., & Skiadopoulos, G. (2008). Measuring the Market Risk of Freight Rates: A Value-At-Risk Approach. International Journal of Theoretical and Applied Finance, 11:5, 447-469.

[9]Barone-Adesi, G., Giannopoulos, K., & Vosper, L. (1999). VaR without Correlations for Nonlinear Portfolios. Journal of Futures Market, 19, 583-602.

[10]Becker, R., & Clements, A. (2008). Are Combination Forecasts of S&P500 Volatility Superior? International Journal of Forecasting, 24, 122-133.

[11]Berkowitz, J. and O'Brien, J. (2002). "How Accurate are Value at Risk Models at Commercial Banks", The Journal of Finance , 57, 2231–2272.

[12]Bessis, J. (2002). Risk Management in Banking 2nd Edition. Wiley.

[13]Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31, 307-327.

[14]Christoffersen, P. (1998). Evaluating Interval Forecast. International Economic Review, 39, 841-862.

[15]Christoffersen, P., Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach. Journal of Financial Econometrics, 2, 84-108.

[16]Colletaz, G., Hurlin, C., & Perignon, C. (2013). The Risk Map: A New Tool for Backtesting Value-at-Risk Models. Journal of Banking and Finance, 37 (10), 3843-3854.

[17]Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. Journal of Financial Econometrics, 7, 174-196.

[18]Corsi, F. & R. Reno (2009). HAR volatility modelling with heterogeneous leverage and jumps, Working paper, University of Siena.

[19]Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation. Econometrica, 50, 987-1008.

[20]Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economic Statistics, 22, 367-381.

[21]Giot, P., and Laurent, S. (2003). Value at risk for long and short trading positions. Journal of Applied Econometrics, 18, 641–664.

[22]Goulas, L., & Skiadopoulos, G. (2012). Are Freight Futures Markets Efficient? Evidence from IMAREX. International Journal of Forecasting, 28, 644-659.

[23]Gurrola-Perez, P. and Murphy, D., (2015). Filtered Historical Simulation VaR Models and their Competitors. Bank of England Working paper No. 525.

[24]Halbleib, R., & Winfried P. (2012). Improving the Value at Risk Forecasts: Theory and Evidence from the Financial Crisis. Journal of Economic Dynamics and Control, 36 (8), 1212-1228 .

[25]Huang, H., & Lee, T., (2010). To combine Forecasts or to combine Information?. Econometric Reviews, 29, 534-570.

[26]Kavussanos, M., & Dimitrakopoulos, D. (2011). Market Risk Model Selection and Medium Term Risk with Limited Data: Application to Ocean Tanker Freight Markets. International Review of Financial Analysis, 20, 258-268.

[27]Kavussanos, M., & Visvikis, I. (2004). Market interactions in returns and volatilities between spot and forward shipping freight derivatives. Journal of Banking and Finance, 28, 2015–2049.

[28]Kavussanos, M. G., & Visvikis, I. D. (2006). Shipping freight derivatives: a survey of recent evidence. Maritime Policy and Management, 33, 233–255.

[29]Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. The American Economic Review 99 (3), 1053–1069.

[30]McAleer, M., Jimenez-Martin, J. A., & Perez-Amaral T. (2013). International Evidence on GFC-Robust Forecasts for Risk Management under the Basel Accord. Journal of Forecasting, 32(3), 267-288.

[31]McNeil, A. J., & Frey, R., (2000). Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. Journal of Empirical Finance, 7, 271–300.

[32]Nomikos N., Kyriakou, I., Papapostolou, N., & Pouliasis P. (2013). Freight Options: Price Modelling and Empirical Analysis. Transportation Research Part E: Logistics and Transportation Review, 51, 82-94.

[33]Nelson, D. (1991). Conditional Heteroscedasticity in Asset Returns: A new approach. Econometrica, 59, 347-370.

[34]O'Brien, J., & Szerszen, P., (2014). An Evaluation of Bank VaR Measures for Market Risk During and Before the Financial Crisis, Working Paper.

[35]Patton, A., Sheppard, K., (2009). Optimal combinations of realized volatility estimators. International Journal of Forecasting, 25, 218-238.

[36]Pritsker, M., (2006). The Hidden Dangers of Historical Simulation. Journal of Banking and Finance, 30, 561-582.

[37]Rapach, D., Strauss, J., & Zhou, G. (2010). Out of Sample Equity Premium Prediction:Combination Forecasts and Links ti the Real Economy. The Review of Financial Studies, 23:2, 821-862.

[38]Rocco, M., (2014). Extreme Value Theory for Finance: A Survey, Journal of Economic Strategy, 28, 82–108.

[39]Sener, E., Baronyan, S., & Menguturk, L. (2012). Ranking the Predictive Performances of Value at Risk estimation methods. International Journal of Forecasting, 28, 849-873.

[40]Stock, J. H., & Watson M.W. (2004). Combination Forecasts of Output Growth in a Seven-Country Data Set. Journal of Forecasting, 23, 405-430.

[41]Stopford, M. (2009). Maritime economics 3rd ed. Routledge, Oxon.

[42]Timmermann, A. (2006). Forecast combinations. In G.Elliott, C.W. J.Granger, & A.G. Timmermann (EBS), Handbook of economic forecasting. Amsterdam : North Holland Press.

[43]Vlaar, P., J. (2000). Value at Risk models for dutch bond portfolios. Journal of Banking and Finance, 24, 1131–1154.

Table 1. Descriptive Statistics

	LBCI	LBCTI	LBDI	LBDTI	LBPI	4TC AVG CAPE	4TC AVG PANAM	TD3	TD5	TD7	TD9
Mean	-0.006	0.000	-0.014	0.004	-0.010	-0.0353	-0.0051	0.0069	0.0093	0.0043	0.0001
Median	0.000	-0.120	0.057	-0.060	0.048	0.0100	0.0483	-0.2432	-0.1363	-0.1528	-0.0881
Maximum	16.502	25.359	13.658	22.951	12.836	31.497	13.112	39.961	55.587	42.700	47.923
Minimum	-19.215	-29.647	-11.953	-38.122	-21.623	-34.203	-21.663	-50.199	-40.180	-49.959	-51.748
Std. Dev.	2.630	1.493	1.832	2.230	2.308	3.679	2.414	5.149	4.776	5.202	6.150
Skewness	0.067	-1.515	-0.066	-1.613	-0.471	0.326	-0.380	0.155	0.694	0.797	0.713
Kurtosis	9.488	93.744	9.928	40.501	12.081	13.675	10.714	15.929	15.968	17.776	16.226
Jarque-Bera	5423.5	1061713.	6184.5	182467.8	10735.7	14731.09	7737.52	21539.86	21905.46	28445.34	22790.52
Probability	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q(1)	1724.3	783.6	2084.1	943.2	2229.0	1734.30	2229.20	807.30	525.07	767.06	417.79
Probability	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q(10)	2767.0	2741.4	4775.0	1974.4	3976.5	2992.50	3997.40	1228.10	726.69	1395.50	608.74
Probability	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q(1)sq	738.38	0.471	1482.0	2.824	1642.4	630.94	1621.70	96.86	25.41	46.26	80.51
Probability	0.000	0.493	0.000	0.093	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q(10)sq	1781.8	2.327	4007.6	7.036	4115.9	1607.00	4013.20	208.61	36.39	103.27	104.48
Probability	0.000	0.993	0.000	0.722	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	3091	3091	3091	3091	3091	3091	3091	3091	3091	3091	3091

Notes: Descriptive statistics are calculated for the whole sample of logarithmic returns covering the period 1/11/1999 -12/03/2012. The Jarque-Bera statistic for the normality test is $\chi^2\,(2)$ distributed. Q(i) and Q(i)sq are the Ljung-Box Q statistics for the returns and squared returns at the i lag.

Table 2. Backtesting results BDI

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.820**	**0.866**	**0.961**	**0.126**	**0.820**	**0.822**	**0.814**	**0.863**	**0.248**	**0.497**	**0.960**	**0.526**	**0.650**	**0.947**	**0.750**
Historical Simulation-250	**0.108**	0.000	0.000	0.003	**0.400**	**0.462**	**0.520**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	0.002	0.000	0.000	0.000	**0.189**	**0.275**	**0.276**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.841**	0.000	0.000	**0.813**	**0.919**	**0.924**	**0.927**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.008	0.000	0.000	0.000	0.000	**0.345**
AR-ARCH(1)-N	0.033	**0.461**	**0.078**	0.000	0.023	0.029	0.040	**0.428**	**0.602**	**0.684**	**0.060**	0.048	**0.116**	**0.138**	**0.697**
AR-ARCH(1)-t	0.000	**0.144**	0.000	0.000	0.000	0.000	0.000	**0.084**	**0.309**	**0.962**	0.000	0.000	0.000	0.000	**0.161**
AR-GARCH(1,1)-N	**0.217**	**0.877**	**0.461**	**0.092**	**0.232**	**0.237**	**0.231**	**0.886**	**0.840**	**0.867**	**0.481**	**0.689**	**0.775**	**0.396**	**0.894**
AR-GARCH(1,1)-t	**0.623**	**0.612**	**0.779**	**0.050**	**0.632**	**0.650**	**0.646**	**0.607**	**0.232**	**0.621**	**0.778**	**0.595**	**0.616**	**0.739**	**0.641**
AR-EGARCH(1,1)-N	**0.350**	**0.428**	**0.472**	**0.422**	**0.349**	**0.364**	**0.358**	**0.498**	**0.977**	**0.726**	**0.522**	**0.618**	**0.700**	**0.570**	**0.956**
AR-EGARCH(1,1)-t	**0.820**	**0.190**	**0.412**	**0.534**	**0.827**	**0.833**	**0.821**	**0.142**	0.013	**0.173**	**0.332**	**0.094**	**0.088**	**0.651**	**0.356**
Monte Carlo Simulation	**0.729**	**0.661**	**0.855**	**0.064**	**0.735**	**0.754**	**0.752**	**0.655**	**0.117**	**0.436**	**0.853**	**0.323**	**0.357**	**0.720**	**0.513**
Combination Mean	**0.432**	0.000	0.000	**0.078**	**0.527**	**0.585**	**0.587**	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001
Combination Median	**0.432**	**0.966**	**0.733**	**0.510**	**0.439**	**0.464**	**0.457**	**0.968**	**0.357**	**0.726**	**0.741**	**0.495**	**0.572**	**0.542**	**0.531**
Combination Trim1	**0.523**	0.000	0.000	**0.179**	**0.604**	**0.653**	**0.656**	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.001
Combination Trim2	**0.278**	0.011	0.021	**0.580**	**0.340**	**0.388**	**0.384**	0.004	0.000	0.002	0.008	0.001	0.002	0.019	0.011

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.269**	**0.534**	**0.447**	0.025	**0.237**	**0.265**	**0.257**	**0.606**	**0.234**	**0.464**	**0.444**	**0.094**	**0.158**	**0.602**	**0.554**
Historical Simulation-250	**0.054**	0.000	0.000	**0.090**	**0.131**	**0.238**	**0.259**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	0.015	0.000	0.000	0.000	**0.242**	**0.263**	**0.326**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	0.000	0.005	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.097**	0.001	0.001	**0.130**	**0.139**	**0.207**	**0.176**	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.003	**0.445**	0.010	0.004	0.001	0.002	0.001	**0.206**	**0.078**	**0.376**	0.002	0.002	0.004	0.027	**0.822**
AR-ARCH(1)-N	0.000	**0.709**	0.000	0.000	0.000	0.000	0.000	**0.501**	**0.161**	**0.064**	0.000	0.000	0.000	0.000	**0.279**
AR-ARCH(1)-t	0.000	**0.210**	0.000	0.000	0.000	0.000	0.000	**0.142**	**0.467**	**0.424**	0.000	0.000	0.000	0.000	**0.548**
AR-GARCH(1,1)-N	0.015	**0.359**	0.033	0.021	0.008	0.011	0.009	**0.131**	0.036	**0.220**	0.008	0.008	0.012	**0.094**	**0.936**
AR-GARCH(1,1)-t	**0.269**	**0.534**	**0.447**	**0.083**	**0.237**	**0.265**	**0.257**	**0.606**	**0.234**	**0.464**	**0.444**	**0.094**	**0.158**	**0.602**	**0.554**
AR-EGARCH(1,1)-N	0.003	**0.445**	0.010	0.007	0.001	0.002	0.001	**0.206**	**0.078**	**0.376**	0.002	0.002	0.004	0.022	**0.516**
AR-EGARCH(1,1)-t	**0.413**	**0.564**	**0.605**	**0.080**	**0.388**	**0.418**	**0.411**	**0.645**	**0.180**	**0.374**	**0.625**	**0.093**	**0.160**	**0.829**	**0.749**
Monte Carlo Simulation	**0.166**	**0.212**	**0.176**	0.026	**0.161**	**0.183**	**0.173**	0.041	0.005	**0.054**	0.042	0.018	0.030	**0.253**	**0.215**
Combination Mean	**0.146**	**0.791**	**0.335**	**0.328**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.184**	**0.860**
Combination Median	**0.930**	**0.658**	**0.903**	**0.387**	**0.929**	**0.938**	**0.940**	**0.756**	**0.068**	**0.164**	**0.949**	0.038	**0.076**	**0.881**	**0.767**
Combination Trim1	**0.146**	**0.791**	**0.335**	**0.398**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.184**	**0.860**
Combination Trim2	**0.682**	**0.690**	**0.849**	**0.655**	**0.686**	**0.685**	**0.718**	**0.791**	**0.706**	**0.135**	**0.891**	**0.960**	0.043	**0.261**	**0.144**

Notes: The Table reports p-values of the employed backtesting criteria. Bold indicates significance at the 5% level and suggests a 'pass' in the respective criterion.

Table 3. Backtesting results Aggregate Indices

Panel A: 5%	BPI			BCI			BDTI			BCTI		
	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC
Filtered Historical Simulation	0	2	0	0	0	0	1	0	0	0	0	0
Historical Simulation-250	1	5	5	0	5	5	2	5	5	2	4	4
Historical Simulation-500	2	5	5	1	5	5	2	5	5	2	4	4
Historical SimulationAll	5	5	5	5	5	5	2	5	5	2	4	4
Riskmetrics	1	5	5	0	5	5	5	5	5	5	3	4
Variance-Covariance	5	5	5	5	5	5	5	5	5	4	4	5
Extreme Value Theory	5	4	5	5	1	5	5	3	5	3	2	4
AR-ARCH(1)-N	0	0	0	5	0	5	5	1	5	3	2	4
AR-ARCH(1)-t	5	0	5	5	0	5	5	0	5	3	2	4
AR-GARCH(1,1)-N	0	1	0	0	0	0	5	3	5	3	2	4
AR-GARCH(1,1)-t	0	1	0	0	0	0	5	0	5	3	0	1
AR-EGARCH(1,1)-N	0	3	0	0	0	0	5	3	5	3	2	4
AR-EGARCH(1,1)-t	0	4	4	0	0	0	4	0	2	3	0	1
Monte Carlo Simulation	0	2	0	0	0	0	5	0	5	3	0	0
Combination Mean	1	5	5	0	4	4	5	0	5	3	2	4
Combination Median	1	1	1	0	0	0	5	1	5	3	2	4
Combination Trim1	2	5	5	0	4	3	5	0	5	3	2	4
Combination Trim2	3	5	5	0	3	2	5	0	5	3	2	4

Panel B: 1%	BPI			BCI			BDTI			BCTI		
	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC
Filtered Historical Simulation	0	0	0	0	0	0	0	0	0	0	0	0
Historical Simulation-250	1	5	5	0	5	5	0	0	0	2	4	4
Historical Simulation-500	1	5	5	0	5	5	0	3	3	2	4	4
Historical SimulationAll	2	5	5	5	5	5	0	3	3	1	3	2
Riskmetrics	2	5	5	0	5	5	0	0	0	0	0	0
Variance-Covariance	5	5	5	5	5	5	0	3	3	2	4	5
Extreme Value Theory	5	0	5	5	0	5	0	0	0	0	0	0
AR-ARCH(1)-N	5	0	5	5	2	5	1	0	0	0	0	1
AR-ARCH(1)-t	1	1	1	5	0	5	0	0	0	0	0	1
AR-GARCH(1,1)-N	2	0	2	5	0	5	0	0	0	0	0	0
AR-GARCH(1,1)-t	0	0	0	0	0	0	1	0	0	0	0	1
AR-EGARCH(1,1)-N	0	0	1	5	0	4	0	1	2	0	0	0
AR-EGARCH(1,1)-t	0	0	0	0	0	0	3	0	1	0	0	1
Monte Carlo Simulation	0	0	0	0	0	0	1	0	0	0	0	1
Combination Mean	0	4	3	0	0	0	1	0	0	0	0	1
Combination Median	0	0	0	0	0	0	1	0	0	0	0	1
Combination Trim1	4	2	3	0	0	0	4	0	2	1	0	1
Combination Trim2	0	0	1	0	0	0	1	0	0	0	0	1

Notes: The table reports the number of 5% and 1% VaR estimation methods failed tests at the 5% significance level.

Table 4. Performance Evaluation: All methods

Panel A: 5%	BDI			BPI			BCI			BDTI			BCTI		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.150	3.614	NR	0.135	3.158	NR	0.276	4.195	NR	0.145	4.753	NR	0.106	5.682	R
Historical Simulation-250	0.313	7.538	R	0.301	7.072	R	0.326	4.956	NR	0.216	7.094	R	0.106	5.674	R
Historical Simulation-500	0.426	10.271	R	0.473	11.101	R	0.440	6.685	R	0.233	7.654	R	0.141	7.548	R
Historical SimulationAll	0.592	14.265	R	0.643	15.091	R	0.936	14.231	R	0.240	7.866	R	0.094	5.058	R
Riskmetrics	0.200	4.822	NR	0.212	4.967	NR	0.287	4.356	NR	0.132	4.325	NR	0.106	5.665	NR
Variance-Covariance	0.489	11.773	R	0.592	13.889	R	0.692	10.513	R	0.221	7.243	R	0.138	7.386	R
Extreme Value Theory	0.208	5.011	NR	0.181	4.238	NR	0.373	5.663	NR	0.289	9.469	R	0.166	8.910	R
AR-ARCH(1)-N	0.210	5.052	NR	0.203	4.769	NR	0.404	6.141	R	0.141	4.616	NR	0.090	4.823	NR
AR-ARCH(1)-t	0.234	5.628	R	0.239	5.611	R	0.472	7.176	R	0.141	4.626	NR	0.097	5.210	NR
AR-GARCH(1,1)-N	0.137	3.291	NR	0.124	2.920	NR	0.260	3.945	NR	0.138	4.523	NR	0.098	5.273	NR
AR-GARCH(1,1)-t	0.139	3.341	NR	0.133	3.117	NR	0.270	4.104	NR	0.142	4.647	NR	0.099	5.286	NR
AR-EGARCH(1,1)-N	0.135	3.255	NR	0.130	3.063	NR	0.248	3.770	NR	0.204	6.676	R	0.089	4.777	NR
AR-EGARCH(1,1)-t	0.147	3.540	NR	0.125	2.923	NR	0.274	4.162	NR	0.144	4.731	NR	0.109	5.826	NR
Monte Carlo Simulation	0.143	3.447	NR	0.132	3.110	NR	0.277	4.209	NR	0.137	4.507	NR	0.101	5.399	NR
Combination Mean	0.167	4.022	NR	0.180	4.215	NR	0.267	4.060	NR	0.130	4.278	NR	0.081	4.331	NR
Combination Median	0.142	3.419	NR	0.129	3.032	NR	0.248	3.769	NR	0.132	4.339	NR	0.079	4.250	NR
Combination Trim1	0.162	3.896	NR	0.169	3.971	NR	0.265	4.030	NR	0.132	4.331	NR	0.084	4.479	NR
Combination Trim2	0.158	3.815	NR	0.160	3.753	NR	0.266	4.036	NR	0.132	4.323	NR	0.082	4.422	NR
Panel B: 1%	BDI			BPI			BCI			BDTI			BCTI		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.023	2.699	NR	0.018	2.014	NR	0.036	2.812	NR	0.021	3.212	NR	0.021	4.273	NR
Historical Simulation-250	0.071	8.318	R	0.048	5.540	NR	0.076	5.899	NR	0.020	3.059	NR	0.072	14.673	R
Historical Simulation-500	0.053	6.234	NR	0.068	7.770	R	0.099	7.690	R	0.080	12.315	R	0.064	13.162	R
Historical SimulationAll	0.193	22.680	R	0.153	17.646	R	0.086	6.697	R	0.080	12.377	R	0.058	11.901	R
Riskmetrics	0.047	5.525	NR	0.036	4.125	NR	0.063	4.912	NR	0.027	4.176	NR	0.021	4.313	NR
Variance-Covariance	0.160	18.772	R	0.251	28.834	R	0.333	26.003	R	0.089	13.683	R	0.060	12.287	R
Extreme Value Theory	0.022	2.549	NR	0.018	2.030	NR	0.035	2.747	NR	0.019	2.844	NR	0.023	4.737	NR
AR-ARCH(1)-N	0.057	6.668	R	0.050	5.767	R	0.158	12.347	R	0.026	3.969	NR	0.017	3.541	NR
AR-ARCH(1)-t	0.032	3.776	NR	0.031	3.573	NR	0.056	4.393	NR	0.023	3.469	NR	0.014	2.778	NR
AR-GARCH(1,1)-N	0.027	3.177	NR	0.023	2.639	NR	0.047	3.624	NR	0.027	4.194	NR	0.022	4.412	NR
AR-GARCH(1,1)-t	0.022	2.634	NR	0.017	1.952	NR	0.035	2.710	NR	0.021	3.162	NR	0.014	2.844	NR
AR-EGARCH(1,1)-N	0.027	3.120	NR	0.022	2.568	NR	0.047	3.665	NR	0.093	14.399	R	0.020	4.011	NR
AR-EGARCH(1,1)-t	0.021	2.417	NR	0.017	1.908	NR	0.034	2.628	NR	0.019	2.981	NR	0.015	3.149	NR
Monte Carlo Simulation	0.022	2.628	NR	0.017	1.936	NR	0.035	2.749	NR	0.021	3.158	NR	0.014	2.831	NR
Combination Mean	0.018	2.079	NR	0.037	4.250	NR	0.037	2.847	NR	0.021	3.287	NR	0.014	2.875	NR
Combination Median	0.021	2.478	NR	0.018	2.064	NR	0.037	2.876	NR	0.022	3.426	NR	0.014	2.845	NR
Combination Trim1	0.017	2.007	NR	0.022	2.573	NR	0.033	2.576	NR	0.020	3.053	NR	0.013	2.648	NR
Combination Trim2	0.019	2.239	NR	0.024	2.810	NR	0.036	2.825	NR	0.021	3.234	NR	0.013	2.720	NR

Notes: The Table reports the Performance Evaluation results for all the implemented methods. The PM column provides the Penalization Measure while the Ratio column provides the corresponding ratio. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Table 5. Performance Evaluation: Second Stage

Panel A: 5%	BDI			BPI			BCI			BDTI			BCTI		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.150	6.437	NR	0.135	5.976	NR	0.276	6.162	NR	0.145	7.834	R	0.106	7.461	R
Historical Simulation-250	-	-	-	-	-	-	0.326	7.280	NR	-	-	-	0.106	7.451	R
Historical Simulation-500	-	-	-	-	-	-	0.440	9.819	R	-	-	-	-	-	-
Historical SimulationAll	-	-	-	-	-	-	0.000	-	-	-	-	-	0.094	6.642	R
Riskmetrics	0.200	8.587	R	0.212	9.400	R	0.287	6.398	R	0.132	7.128	NR	0.106	7.439	R
Variance-Covariance	-	-	-	-	-	-	0.000	-	-	-	-	-	-	-	-
Extreme Value Theory	0.208	8.923	R	0.181	8.019	R	0.373	8.318	R	-	-	-	-	-	-
AR-ARCH(1)-N	0.210	8.997	R	0.203	9.025	R	0.404	9.020	R	0.141	7.607	NR	0.090	6.333	NR
AR-ARCH(1)-t	0.234	10.023	R	0.239	10.618	R	0.000	-	-	0.141	7.624	R	0.097	6.841	R
AR-GARCH(1,1)-N	0.137	5.861	NR	0.124	5.525	NR	0.260	5.795	NR	0.138	7.455	NR	0.098	6.924	NR
AR-GARCH(1,1)-t	0.139	5.950	NR	0.133	5.898	NR	0.270	6.028	NR	0.142	7.659	R	0.099	6.941	R
AR-EGARCH(1,1)-N	0.135	5.797	NR	0.130	5.795	NR	0.248	5.537	NR	0.204	11.002	R	0.089	6.272	NR
AR-EGARCH(1,1)-t	0.147	6.304	NR	0.125	5.530	NR	0.274	6.113	NR	0.144	7.798	R	0.109	7.650	R
Monte Carlo Simulation	0.143	6.138	NR	0.132	5.884	NR	0.277	6.182	NR	0.137	7.428	NR	0.101	7.090	R
Combination Mean	0.167	7.162	NR	0.180	7.976	R	0.267	5.964	NR	0.130	7.050	NR	0.081	5.688	NR
Combination Median	0.142	6.089	NR	0.129	5.737	NR	0.248	5.536	NR	0.132	7.152	NR	0.079	5.581	NR
Combination Trim1	0.162	6.938	NR	0.169	7.514	R	0.265	5.920	NR	0.132	7.138	NR	0.084	5.882	NR
Combination Trim2	0.158	6.795	NR	0.160	7.102	NR	0.266	5.929	NR	0.132	7.124	NR	0.082	5.806	NR
Panel B: 1%	BDI			BPI			BCI			BDTI			BCTI		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.023	6.195	NR	0.018	5.037	NR	0.036	5.211	NR	0.021	6.802	NR	0.021	8.906	R
Historical Simulation-250	-	-	-	0.048	13.857	R	0.076	10.933	R	0.020	6.478	R	-	-	-
Historical Simulation-500	0.053	14.310	R	-	-	-	-	-	-	-	-	-	-	-	-
Historical SimulationAll	-	-	-	-	-	-	0.086	12.412	R	-	-	-	-	-	-
Riskmetrics	0.047	12.683	R	0.036	10.317	R	0.063	9.103	R	0.027	8.842	R	0.021	8.989	R
Variance-Covariance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Extreme Value Theory	0.022	5.851	NR	0.018	5.076	NR	0.035	5.090	NR	0.019	6.022	NR	0.023	9.875	R
AR-ARCH(1)-N	-	-	-	-	-	-	-	-	-	0.026	8.404	R	0.017	7.380	R
AR-ARCH(1)-t	0.032	8.668	R	0.031	8.936	R	0.056	8.141	R	0.023	7.346	NR	0.014	5.791	NR
AR-GARCH(1,1)-N	0.027	7.294	R	0.023	6.601	NR	0.047	6.716	R	0.027	8.880	R	0.022	9.196	R
AR-GARCH(1,1)-t	0.022	6.047	NR	0.017	4.882	NR	0.035	5.022	NR	0.021	6.696	NR	0.014	5.927	NR
AR-EGARCH(1,1)-N	0.027	7.162	R	0.022	6.423	NR	0.047	6.791	R	-	-	-	0.020	8.360	R
AR-EGARCH(1,1)-t	0.021	5.548	NR	0.017	4.773	NR	0.034	4.871	NR	0.019	6.312	NR	0.015	6.563	NR
Monte Carlo Simulation	0.022	6.032	NR	0.017	4.842	NR	0.035	5.095	NR	0.021	6.688	NR	0.014	5.901	NR
Combination Mean	0.018	4.772	NR	0.037	10.631	R	0.037	5.277	NR	0.021	6.961	NR	0.014	5.993	NR
Combination Median	0.021	5.688	NR	0.018	5.162	NR	0.037	5.330	NR	0.022	7.255	R	0.014	5.931	NR
Combination Trim1	0.017	4.608	NR	0.022	6.435	NR	0.033	4.773	NR	0.020	6.465	NR	0.013	5.519	NR
Combination Trim2	0.019	5.141	NR	0.024	7.028	R	0.036	5.235	NR	0.021	6.849	NR	0.013	5.670	NR

Notes: The Table reports the second stage Performance Evaluation results. For each index the worst performing methods are excluded and the analysis is repeated with the remaining methods. The PM column provides the Penalization Measure while the Ratio column provides the corresponding ratio. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Table 6. Ranking and Performance Equality Test

Panel A: 5%

	BDI		BPI		BCI		BDTI		BCTI	
1	AR-EGARCH(1,1)-N	NR	AR-GARCH(1,1)-N	NR	Combination Median	NR	Combination Mean	NR	Combination Median	NR
2	AR-GARCH(1,1)-N	NR	Combination Median	NR	AR-EGARCH(1,1)-N	NR	Combination Trim2	NR	Combination Mean	NR
3	AR-GARCH(1,1)-t	NR	AR-EGARCH(1,1)-N	NR	AR-GARCH(1,1)-N	NR	Riskmetrics	NR	Combination Trim2	NR
4	Combination Median	NR	Monte Carlo Simulation	NR	Combination Trim1	NR	Combination Trim1	NR	Combination Trim1	NR
5	Monte Carlo Simulation	NR	AR-GARCH(1,1)-t	NR	Combination Trim2	NR	Combination Median	NR	AR-EGARCH(1,1)-N	NR
6	AR-EGARCH(1,1)-t	NR	Filtered Historical Simulation	NR	AR-GARCH(1,1)-t	NR	Monte Carlo Simulation	NR	AR-ARCH(1)-N	NR
7	Filtered Historical Simulation	NR	AR-ARCH(1)-N	R	AR-EGARCH(1,1)-t	NR	AR-GARCH(1,1)-N	NR	Historical SimulationAll	R
8	-	-	-	-	Filtered Historical Simulation	NR	AR-ARCH(1)-N	NR	AR-ARCH(1)-t	R
9	-	-	-	-	Monte Carlo Simulation	NR	AR-ARCH(1)-t	R	AR-GARCH(1,1)-N	NR
10	-	-	-	-	-	-	AR-GARCH(1,1)-t	R	AR-GARCH(1,1)-t	R
11	-	-	-	-	-	-	AR-EGARCH(1,1)-t	R	Monte Carlo Simulation	R
12	-	-	-	-	-	-	Filtered Historical Simulation	R	Riskmetrics	R
13	-	-	-	-	-	-	AR-EGARCH(1,1)-N	R	Historical Simulation-250	R
14	-	-	-	-	-	-	-	-	Filtered Historical Simulation	R
15	-	-	-	-	-	-	-	-	AR-EGARCH(1,1)-t	R

Panel B: 1%

	BDI		BPI		BCI		BDTI		BCTI	
1	Combination Trim1	NR	AR-EGARCH(1,1)-t	NR	Combination Trim1	NR	Extreme Value Theory	NR	Combination Trim1	NR
2	Combination Mean	NR	Monte Carlo Simulation	NR	AR-EGARCH(1,1)-t	NR	AR-EGARCH(1,1)-t	NR	Combination Trim2	NR
3	Combination Trim2	NR	AR-GARCH(1,1)-t	NR	AR-GARCH(1,1)-t	NR	Combination Trim1	NR	AR-ARCH(1)-t	NR
4	AR-EGARCH(1,1)-t	NR	Filtered Historical Simulation	NR	Monte Carlo Simulation	NR	Historical Simulation-250	R	Monte Carlo Simulation	NR
5	Combination Median	NR	Combination Median	NR	Filtered Historical Simulation	NR	Monte Carlo Simulation	NR	AR-GARCH(1,1)-t	NR
6	Monte Carlo Simulation	NR	AR-EGARCH(1,1)-N	NR	Combination Trim2	NR	AR-GARCH(1,1)-t	NR	Combination Median	NR
7	AR-GARCH(1,1)-t	NR	AR-GARCH(1,1)-N	NR	Combination Mean	NR	Filtered Historical Simulation	NR	Combination Mean	NR
8	Filtered Historical Simulation	NR	Combination Trim2	R	Combination Median	NR	Combination Trim2	NR	AR-EGARCH(1,1)-t	NR
9	-	-	AR-ARCH(1)-t	R	-	-	Combination Mean	NR	AR-ARCH(1)-N	R
10	-	-	Combination Mean	R	-	-	Combination Median	R	AR-EGARCH(1,1)-N	R
11	-	-	-	-	-	-	AR-ARCH(1)-t	NR	Filtered Historical Simulation	R
12	-	-	-	-	-	-	AR-ARCH(1)-N	R	Riskmetrics	R
13	-	-	-	-	-	-	Riskmetrics	R	AR-GARCH(1,1)-N	R
14	-	-	-	-	-	-	AR-GARCH(1,1)-N	R	Extreme Value Theory	R

Notes: The Table reports the ranking of the statistically accepted methods according to the PM measure presented in Table 8. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Table 7. Backtesting results Individual Routes

Panel A: 5%

	4 TC Av Ca			4 TC Av Pa			TD3			TD5			TD7			TD9		
	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC
Filtered Historical Simulation	5	1	4	0	1	0	0	4	4	0	1	0	0	0	0	0	0	0
Historical Simulation-250	1	5	5	1	5	5	0	5	5	0	1	0	1	3	5	1	3	5
Historical Simulation-500	3	5	5	2	5	5	0	5	5	0	1	0	2	4	5	1	3	5
Historical SimulationAll	5	5	5	5	5	5	1	5	5	0	4	4	1	3	5	1	3	5
Riskmetrics	0	5	5	1	5	5	0	5	5	0	1	2	4	3	3	2	4	4
Variance-Covariance	5	5	5	4	5	5	5	5	5	0	2	1	2	4	5	1	3	5
Extreme Value Theory	5	4	5	5	4	5	5	3	5	5	2	5	4	5	5	4	3	4
AR-ARCH(1)-N	5	0	5	0	0	0	5	0	5	0	1	0	4	3	3	0	1	0
AR-ARCH(1)-t	5	0	5	5	0	1	0	0	2	5	0	5	0	0	0	4	3	4
AR-GARCH(1,1)-N	1	2	0	0	0	0	5	2	5	0	1	1	4	3	3	1	0	0
AR-GARCH(1,1)-t	1	1	1	0	0	0	0	0	0	5	0	5	0	0	0	4	3	3
AR-EGARCH(1,1)-N	1	1	0	0	4	4	5	4	5	0	1	0	4	3	3	0	0	0
AR-EGARCH(1,1)-t	5	0	3	0	4	4	0	0	0	0	1	1	0	0	1	4	3	3
Monte Carlo Simulation	1	0	0	0	0	0	0	0	0	5	0	5	0	0	0	4	3	3
Combination Mean	0	4	4	3	5	5	3	4	5	0	0	0	1	0	1	0	0	0
Combination Median	0	0	0	1	4	5	5	2	5	0	2	0	3	3	2	0	0	0
Combination Trim1	0	4	4	3	5	5	3	4	5	0	0	0	1	0	1	0	0	0
Combination Trim2	0	4	4	3	5	5	3	4	5	0	0	0	1	0	1	0	0	0

Panel B: 1%

	4 TC Av Ca			4 TC Av Pa			TD3			TD5			TD7			TD9		
	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC	UC	IND	CC
Filtered Historical Simulation	4	2	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Historical Simulation-250	1	5	5	2	5	5	0	4	4	0	3	3	0	2	4	0	1	1
Historical Simulation-500	3	5	5	1	5	5	0	4	4	0	1	2	0	0	0	0	3	4
Historical SimulationAll	5	5	5	2	5	5	0	4	4	0	0	0	0	0	2	0	3	4
Riskmetrics	3	5	5	2	5	5	5	4	5	5	1	4	0	1	0	4	4	4
Variance-Covariance	5	5	5	5	5	5	0	4	4	5	4	5	4	5	4	4	5	4
Extreme Value Theory	5	2	5	5	0	4	0	0	0	0	0	0	0	0	0	0	0	0
AR-ARCH(1)-N	5	1	5	5	0	5	0	0	0	5	2	4	3	3	2	4	2	3
AR-ARCH(1)-t	5	0	5	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0
AR-GARCH(1,1)-N	5	2	5	2	0	2	0	3	3	0	2	2	0	0	0	3	1	0
AR-GARCH(1,1)-t	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
AR-EGARCH(1,1)-N	5	2	5	0	0	1	0	3	3	0	0	0	0	0	2	3	1	0
AR-EGARCH(1,1)-t	5	0	4	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
Monte Carlo Simulation	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Combination Mean	0	3	2	0	4	3	3	0	1	0	0	0	0	0	2	0	0	0
Combination Median	5	2	5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Combination Trim1	0	1	2	1	5	5	3	0	1	0	0	0	0	0	0	0	0	0
Combination Trim2	1	1	2	0	4	3	1	0	0	0	0	0	0	0	2	0	0	0

Notes: The table reports the number of 5% and 1% VaR estimation methods failed tests at the 5% significance level.

Table 8. Ranking and Performance Equality Test

	4 TC CAPE		4 TC PAN		TD3	
Panel A: 5%						
1	AR-EGARCH11-Normal Distibution	NR	AR-GARCH11-Normal Distibution	NR	AR-EGARCH11-Student Distribution	NR
2	AR-GARCH11-Normal Distibution	NR	Montecarlo Simulation	NR	Montecarlo Simulation	R
3	Combination Median	NR	AR-GARCH11-Student Distribution	NR	AR-ARCH1-Student Distribution	R
4	AR-GARCH11-Student Distribution	NR	Filtered Historical Simulation	NR	AR-GARCH11-Student Distribution	R
5	Montecarlo Simulation	NR	-		-	
6	AR-EGARCH11-Student Distribution	NR	-		-	
7	-		-		-	
8	-		-		-	
9	-		-		-	
10	-		-		-	
11	-		-		-	
12	-		-		-	
13	-		-		-	
14	-		-		-	
Panel B: 1%						
	4 TC CAPE		4 TC PAN		TD3	
1	Combination Trim1	NR	AR-EGARCH11-Student Distribution	NR	Filtered Historical Simulation	NR
2	Montecarlo Simulation	NR	Montecarlo Simulation	NR	AR-ARCH1-Student Distribution	NR
3	AR-GARCH11-Student Distribution	NR	AR-GARCH11-Student Distribution	NR	Combination Trim1	NR
4	Combination Mean	NR	Filtered Historical Simulation	NR	Combination Mean	NR
5	Combination Trim2	NR	Combination Median	NR	Combination Trim2	NR
6	-		AR-EGARCH11-Normal Distibution	NR	Combination Median	NR
7	-		AR-GARCH11-Normal Distibution	NR	Extreme Value Theory	R
8	-		AR-ARCH1-Student Distribution	R	Historical Simulation -Two Year Data	R
9	-		Combination Trim2	R	AR-EGARCH11-Student Distribution	R
10	-		Combination Mean	R	AR-GARCH11-Student Distribution	R
11	-		Historical Simulation -One Year Data	R	Montecarlo Simulation	R
12	-		-		-	
13	-		-		-	
14	-		-		-	

Notes: The Table reports the ranking of the statistically accepted methods according to the PM measure presented in Table A9. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Table 8.(cont'd) Ranking and Performance Equality Test

Panel A: 5%						
	TD5		TD7		TD9	
1	Riskmetrics	NR	Combination Median	NR	AR-EGARCH11-Normal Distibution	NR
2	Filtered Historical Simulation	R	Combination Mean	NR	AR-GARCH11-Normal Distibution	NR
3	AR-EGARCH11-Student Distribution	R	Combination Trim1	NR	Combination Median	NR
4	AR-EGARCH11-Normal Distibution	NR	Combination Trim2	NR	Combination Mean	NR
5	Combination Median	NR	Montecarlo Simulation	R	Combination Trim1	NR
6	Combination Mean	NR	Filtered Historical Simulation	R	AR-ARCH1-Normal Distribution	NR
7	AR-GARCH11-Normal Distibution	NR	AR-EGARCH11-Student Distribution	R	Combination Trim2	NR
8	Combination Trim1	R	AR-GARCH11-Student Distribution	R	Filtered Historical Simulation	NR
9	Combination Trim2	R	-		-	
10	Historical Simulation -One Year Data	R	-		-	
11	AR-ARCH1-Normal Distribution	R	-		-	
12	Historical Simulation -Two Year Data	R	-		-	
13	Variance-Covariance	R	-		-	
14	-		-		-	
Panel B: 1%						
	TD5		TD7		TD9	
1	Extreme Value Theory	NR	Filtered Historical Simulation	NR	AR-EGARCH11-Student Distribution	NR
2	AR-EGARCH11-Student Distribution	NR	Extreme Value Theory	NR	Extreme Value Theory	NR
3	AR-GARCH11-Student Distribution	NR	AR-GARCH11-Student Distribution	NR	Montecarlo Simulation	NR
4	Combination Trim1	NR	Combination Trim1	NR	AR-ARCH1-Student Distribution	NR
5	Filtered Historical Simulation	NR	Historical Simulation -Two Year Data	NR	Filtered Historical Simulation	NR
6	Montecarlo Simulation	NR	Montecarlo Simulation	NR	AR-GARCH11-Student Distribution	NR
7	Combination Mean	NR	AR-EGARCH11-Student Distribution	R	Combination Trim1	NR
8	Combination Trim2	NR	Combination Median	R	Combination Trim2	NR
9	Combination Median	NR	Riskmetrics	R	Combination Mean	NR
10	AR-ARCH1-Student Distribution	R	Combination Mean	R	Combination Median	R
11	Riskmetrics	R	Combination Trim2	R	AR-GARCH11-Normal Distibution	R
12	Historical Simulation All Sample	NR	Historical Simulation All Sample	R	AR-EGARCH11-Normal Distibution	R
13	Historical Simulation -One Year Data	R	AR-GARCH11-Normal Distibution	R	-	
14	Historical Simulation -Two Year Data	NR	AR-ARCH1-Student Distribution	R	-	

Notes: The Table reports the ranking of the statistically accepted methods according to the PM measure presented in Table A9. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Appendix B. Detailed Tables

Table B1. Backtesting results BPI

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.712**	**0.080**	**0.201**	**0.843**	**0.727**	**0.732**	**0.727**	0.042	0.019	**0.055**	**0.119**	**0.119**	**0.211**	**0.826**	**0.564**
Historical Simulation-250	**0.082**	0.000	0.000	0.000	**0.589**	**0.571**	**0.626**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	0.002	0.000	0.000	0.000	**0.226**	**0.249**	**0.257**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.712**	0.000	0.000	0.036	**0.839**	**0.838**	**0.834**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.013	0.015	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.433**
AR-ARCH(1)-N	**0.354**	**0.872**	**0.643**	**0.350**	**0.343**	**0.345**	**0.376**	**0.866**	**0.844**	**0.137**	**0.631**	**0.768**	**0.165**	**0.186**	**0.096**
AR-ARCH(1)-t	0.000	**0.369**	0.001	0.000	0.000	0.000	0.000	**0.300**	**0.655**	**0.295**	0.000	0.000	0.001	0.002	**0.319**
AR-GARCH(1,1)-N	**0.432**	**0.065**	**0.134**	**0.771**	**0.471**	**0.472**	**0.486**	0.042	**0.172**	**0.134**	**0.093**	**0.190**	**0.269**	**0.640**	**0.804**
AR-GARCH(1,1)-t	**0.932**	**0.166**	**0.381**	**0.871**	**0.935**	**0.933**	**0.928**	**0.122**	0.033	**0.067**	**0.301**	**0.206**	**0.334**	**0.696**	**0.396**
AR-EGARCH(1,1)-N	**0.841**	0.039	**0.116**	**0.813**	**0.853**	**0.859**	**0.868**	0.018	**0.061**	0.035	**0.059**	**0.127**	**0.153**	**0.570**	**0.315**
AR-EGARCH(1,1)-t	**0.932**	0.004	0.016	**0.871**	**0.939**	**0.934**	**0.929**	0.000	0.012	0.027	0.002	0.006	0.015	**0.401**	**0.177**
Monte Carlo Simulation	**0.820**	**0.067**	**0.183**	**0.716**	**0.830**	**0.829**	**0.824**	0.034	0.045	**0.088**	**0.104**	**0.173**	**0.289**	**0.620**	**0.330**
Combination Mean	0.030	0.000	0.000	**0.067**	**0.084**	**0.125**	**0.169**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Combination Median	0.045	**0.220**	**0.063**	**0.113**	**0.067**	**0.081**	**0.096**	**0.226**	**0.094**	0.045	**0.076**	**0.091**	**0.105**	0.019	**0.064**
Combination Trim1	0.013	0.001	0.000	**0.053**	0.045	**0.071**	**0.113**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Combination Trim2	0.013	0.001	0.000	0.010	0.045	**0.071**	**0.101**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.413**	**0.564**	**0.605**	**0.641**	**0.388**	**0.380**	**0.410**	**0.645**	**0.511**	**0.407**	**0.625**	**0.761**	**0.158**	**0.832**	**0.759**
Historical Simulation-250	**0.097**	0.000	0.000	0.007	**0.306**	**0.380**	**0.344**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	**0.054**	0.000	0.000	0.002	**0.314**	**0.282**	**0.333**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	**0.098**	**0.054**	**0.083**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.007	0.000	0.000	0.017	**0.063**	**0.074**	**0.092**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.007	**0.373**	0.018	0.021	0.002	0.003	0.004	**0.377**	**0.720**	**0.266**	0.007	0.008	0.007	0.044	**0.554**
AR-ARCH(1)-N	0.000	**0.264**	0.000	0.000	0.000	0.000	0.000	**0.216**	**0.752**	**0.825**	0.000	0.000	0.000	0.001	**0.240**
AR-ARCH(1)-t	**0.097**	**0.477**	**0.196**	0.024	**0.069**	**0.064**	**0.060**	**0.528**	**0.368**	**0.266**	**0.164**	**0.257**	**0.345**	0.037	0.030
AR-GARCH(1,1)-N	**0.054**	**0.450**	**0.118**	**0.125**	0.032	0.040	**0.050**	**0.490**	**0.443**	**0.097**	**0.086**	0.046	0.023	**0.254**	**0.873**
AR-GARCH(1,1)-t	**0.930**	**0.658**	**0.903**	**0.664**	**0.929**	**0.931**	**0.940**	**0.756**	**0.659**	**0.188**	**0.949**	**0.977**	**0.076**	**0.899**	**0.827**
AR-EGARCH(1,1)-N	**0.097**	**0.477**	**0.196**	**0.197**	**0.069**	**0.083**	**0.099**	**0.528**	**0.365**	**0.064**	**0.164**	**0.066**	0.026	**0.194**	**0.240**
AR-EGARCH(1,1)-t	**0.599**	**0.594**	**0.756**	**0.768**	**0.585**	**0.579**	**0.604**	**0.683**	**0.560**	**0.324**	**0.795**	**0.889**	**0.142**	**0.972**	**0.927**
Monte Carlo Simulation	**0.930**	**0.658**	**0.903**	**0.998**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.625**	**0.379**
Combination Mean	**0.457**	0.048	**0.107**	**0.679**	**0.529**	**0.575**	**0.613**	0.001	0.000	0.000	0.003	0.000	0.000	**0.081**	**0.050**
Combination Median	**0.457**	**0.723**	**0.712**	**0.530**	**0.471**	**0.470**	**0.468**	**0.824**	**0.752**	**0.698**	**0.756**	**0.895**	**0.957**	**0.554**	**0.975**
Combination Trim1	0.007	**0.895**	0.027	0.037	0.023	0.024	**0.135**	**0.959**	**0.943**	0.001	**0.074**	**0.157**	0.000	0.001	0.025
Combination Trim2	**0.457**	**0.723**	**0.712**	**0.679**	**0.471**	**0.470**	**0.526**	**0.824**	**0.752**	**0.092**	**0.756**	**0.895**	0.019	**0.290**	**0.255**

Notes: The Table reports p-values of the employed backtesting criteria. Bold indicates significance at the 5% level and suggests a ‘pass’ in the respective criterion.

Table B2. Backtesting results BCI

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.431**	**0.630**	**0.653**	**0.158**	**0.430**	**0.433**	**0.436**	**0.600**	**0.463**	**0.374**	**0.632**	**0.755**	**0.834**	**0.818**	**0.967**
Historical Simulation-250	**0.516**	0.000	0.000	**0.087**	**0.682**	**0.688**	**0.685**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	**0.061**	0.000	0.000	0.001	**0.251**	**0.275**	**0.288**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.278**	0.000	0.000	**0.469**	**0.497**	**0.544**	**0.511**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	**0.150**	0.000	0.000	0.000	0.000	0.000	0.047	**0.090**	**0.285**	0.000	0.000	0.000	0.000	**0.528**
AR-ARCH(1)-N	0.004	**0.562**	0.012	0.000	0.002	0.003	0.003	**0.510**	**0.519**	**0.618**	0.007	0.006	0.014	0.023	**0.646**
AR-ARCH(1)-t	0.000	**0.487**	0.000	0.000	0.000	0.000	0.000	**0.386**	**0.834**	**0.571**	0.000	0.000	0.000	0.000	**0.800**
AR-GARCH(1,1)-N	**0.278**	**0.930**	**0.553**	**0.219**	**0.291**	**0.308**	**0.324**	**0.934**	**0.539**	**0.343**	**0.569**	**0.629**	**0.675**	**0.481**	**0.925**
AR-GARCH(1,1)-t	**0.610**	**0.972**	**0.878**	**0.149**	**0.607**	**0.602**	**0.604**	**0.971**	**0.961**	**0.672**	**0.875**	**0.965**	**0.948**	**0.928**	**0.886**
AR-EGARCH(1,1)-N	**0.841**	**0.329**	**0.609**	**0.429**	**0.846**	**0.850**	**0.860**	**0.297**	**0.600**	**0.300**	**0.569**	**0.747**	**0.672**	**0.944**	**0.971**
AR-EGARCH(1,1)-t	**0.181**	**0.219**	**0.192**	**0.131**	**0.188**	**0.188**	**0.197**	**0.149**	**0.248**	**0.158**	**0.138**	**0.260**	**0.320**	**0.444**	**0.609**
Monte Carlo Simulation	**0.354**	**0.343**	**0.416**	**0.092**	**0.361**	**0.363**	**0.365**	**0.284**	**0.290**	**0.268**	**0.361**	**0.527**	**0.661**	**0.724**	**0.811**
Combination Mean	**0.165**	0.030	0.037	**0.434**	**0.217**	**0.256**	**0.281**	0.018	0.001	0.001	0.024	0.007	0.008	**0.114**	**0.148**
Combination Median	**0.623**	**0.091**	**0.213**	**0.681**	**0.648**	**0.656**	**0.666**	**0.061**	**0.102**	**0.107**	**0.154**	**0.274**	**0.395**	**0.732**	**0.626**
Combination Trim1	**0.217**	0.037	**0.053**	**0.526**	**0.268**	**0.307**	**0.331**	0.022	0.002	0.001	0.035	0.011	0.013	**0.172**	**0.195**
Combination Trim2	**0.350**	**0.054**	**0.101**	**0.675**	**0.395**	**0.430**	**0.451**	0.034	0.004	0.002	**0.070**	0.027	0.033	**0.273**	**0.225**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.166**	**0.505**	**0.307**	**0.077**	**0.134**	**0.127**	**0.147**	**0.567**	**0.414**	**0.600**	**0.284**	**0.413**	**0.137**	**0.534**	**0.830**
Historical Simulation-250	**0.269**	0.012	0.023	**0.207**	**0.311**	**0.410**	**0.415**	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.006
Historical Simulation-500	**0.166**	0.001	0.001	**0.335**	**0.230**	**0.353**	**0.338**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.269**	0.000	0.001	**0.406**	**0.353**	**0.361**	**0.381**	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.002
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.001	**0.491**	0.005	0.000	0.000	0.000	0.000	**0.252**	**0.931**	**0.444**	0.001	0.001	0.002	0.014	**0.928**
AR-ARCH(1)-N	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.002	**0.334**	**0.334**	0.000	0.000	0.000	0.000	**0.508**
AR-ARCH(1)-t	0.000	**0.227**	0.000	0.000	0.000	0.000	0.000	**0.165**	**0.555**	**0.891**	0.000	0.000	0.000	0.000	**0.927**
AR-GARCH(1,1)-N	0.007	**0.401**	0.019	0.003	0.003	0.003	0.004	**0.166**	**0.727**	**0.263**	0.004	0.007	0.007	0.046	**0.605**
AR-GARCH(1,1)-t	**0.599**	**0.594**	**0.756**	**0.213**	**0.585**	**0.579**	**0.572**	**0.683**	**0.560**	**0.473**	**0.795**	**0.889**	**0.937**	**0.277**	**0.113**
AR-EGARCH(1,1)-N	0.015	**0.398**	0.035	0.004	0.005	0.005	0.006	**0.414**	**0.243**	**0.939**	0.017	0.031	0.026	**0.088**	**0.723**
AR-EGARCH(1,1)-t	**0.269**	**0.534**	**0.447**	**0.083**	**0.237**	**0.230**	**0.222**	**0.606**	**0.462**	**0.364**	**0.444**	**0.592**	**0.699**	**0.540**	**0.450**
Monte Carlo Simulation	**0.413**	**0.564**	**0.605**	**0.452**	**0.388**	**0.380**	**0.372**	**0.645**	**0.511**	**0.417**	**0.625**	**0.761**	**0.845**	**0.358**	**0.182**
Combination Mean	**0.823**	**0.626**	**0.866**	**0.827**	**0.820**	**0.815**	**0.810**	**0.720**	**0.610**	**0.529**	**0.914**	**0.958**	**0.978**	**0.882**	**0.622**
Combination Median	**0.269**	**0.534**	**0.447**	**0.485**	**0.237**	**0.230**	**0.257**	**0.606**	**0.462**	**0.499**	**0.444**	**0.592**	**0.156**	**0.702**	**0.833**
Combination Trim1	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.760**	**0.997**
Combination Trim2	**0.599**	**0.594**	**0.756**	**0.775**	**0.585**	**0.579**	**0.572**	**0.683**	**0.560**	**0.473**	**0.795**	**0.889**	**0.937**	**0.950**	**0.819**

Notes: See Table B1.

Table B3. Backtesting results BDTI

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	0.045	**0.220**	**0.063**	**0.158**	**0.067**	**0.073**	**0.064**	**0.226**	**0.251**	**0.745**	**0.076**	**0.153**	**0.167**	**0.104**	**0.815**
Historical Simulation-250	0.030	0.000	0.000	0.036	**0.120**	**0.166**	**0.203**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	0.008	0.000	0.000	0.017	**0.070**	**0.098**	**0.111**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.003	0.000	0.000	0.017	**0.050**	**0.074**	**0.086**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.000	0.000	0.000	0.000	0.004	0.010	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005
Variance-Covariance	0.000	0.000	0.000	0.000	0.011	0.022	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	**0.120**	0.000	0.000	0.000	0.000	0.000	0.033	0.003	0.002	0.000	0.000	0.000	0.000	**0.438**
AR-ARCH(1)-N	0.000	**0.211**	0.000	0.000	0.000	0.000	0.000	**0.351**	**0.644**	**0.806**	0.000	0.000	0.000	0.000	0.021
AR-ARCH(1)-t	0.000	**0.789**	0.001	0.002	0.001	0.001	0.002	**0.835**	**0.770**	**0.722**	0.004	0.010	0.024	0.000	**0.230**
AR-GARCH(1,1)-N	0.000	0.012	0.000	0.000	0.000	0.000	0.000	0.025	**0.184**	**0.306**	0.000	0.000	0.000	0.000	0.010
AR-GARCH(1,1)-t	0.001	**0.252**	0.001	0.003	0.002	0.003	0.000	**0.307**	**0.415**	**0.793**	0.003	0.010	0.019	0.001	**0.348**
AR-EGARCH(1,1)-N	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.043	**0.052**	**0.166**	0.000	0.000	0.000	0.000	0.018
AR-EGARCH(1,1)-t	0.013	**0.482**	0.036	**0.060**	0.022	0.022	0.019	**0.524**	**0.735**	**0.794**	**0.052**	**0.116**	**0.149**	0.033	**0.722**
Monte Carlo Simulation	0.000	**0.789**	0.001	0.002	0.001	0.001	0.001	**0.835**	**0.744**	**0.514**	0.004	0.009	0.017	0.001	**0.687**
Combination Mean	0.000	**0.359**	0.000	0.000	0.000	0.000	0.000	**0.491**	**0.090**	**0.082**	0.000	0.000	0.000	0.000	**0.134**
Combination Median	0.000	**0.058**	0.000	0.000	0.000	0.000	0.000	**0.095**	**0.118**	**0.303**	0.000	0.000	0.000	0.000	0.031
Combination Trim1	0.000	**0.401**	0.000	0.000	0.000	0.000	0.000	**0.527**	**0.109**	**0.102**	0.000	0.000	0.000	0.000	**0.100**
Combination Trim2	0.000	**0.445**	0.000	0.000	0.000	0.000	0.000	**0.563**	**0.131**	**0.124**	0.000	0.000	0.000	0.000	**0.177**
Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.146**	**0.791**	**0.335**	**0.127**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.173**	**0.695**
Historical Simulation-250	**0.275**	**0.757**	**0.526**	**0.272**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.279**	**0.501**
Historical Simulation-500	**0.275**	**0.757**	**0.526**	**0.272**	**0.301**	**0.458**	**0.458**	**0.856**	0.000	0.000	**0.580**	0.000	0.000	0.048	0.046
Historical SimulationAll	**0.275**	**0.757**	**0.526**	**0.272**	**0.301**	**0.458**	**0.458**	**0.856**	0.000	0.000	**0.580**	0.000	0.000	0.048	0.046
Riskmetrics	**0.275**	**0.757**	**0.526**	**0.272**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.129**	**0.157**
Variance-Covariance	**0.275**	**0.757**	**0.526**	**0.272**	**0.301**	**0.458**	**0.458**	**0.856**	0.000	0.000	**0.580**	0.000	0.000	0.048	0.046
Extreme Value Theory	**0.680**	**0.690**	**0.848**	**0.680**	**0.683**	**0.683**	**0.683**	**0.791**	**0.706**	**0.643**	**0.890**	**0.959**	**0.985**	**0.478**	**0.336**
AR-ARCH(1)-N	**0.066**	**0.825**	**0.179**	0.007	**0.097**	**0.096**	**0.096**	**0.913**	**0.877**	**0.849**	**0.252**	**0.430**	**0.597**	**0.052**	**0.344**
AR-ARCH(1)-t	**0.146**	**0.791**	**0.335**	**0.127**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.187**	**0.982**
AR-GARCH(1,1)-N	**0.146**	**0.791**	**0.335**	**0.127**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.173**	**0.695**
AR-GARCH(1,1)-t	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.052**	**0.345**
AR-EGARCH(1,1)-N	**0.457**	**0.723**	**0.712**	**0.464**	**0.471**	**0.527**	**0.526**	**0.824**	0.027	**0.075**	**0.756**	0.008	0.019	**0.279**	**0.241**
AR-EGARCH(1,1)-t	0.025	**0.860**	**0.079**	0.010	0.049	**0.050**	**0.050**	**0.938**	**0.911**	**0.891**	**0.144**	**0.275**	**0.423**	0.011	**0.168**
Monte Carlo Simulation	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.052**	**0.345**
Combination Mean	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.052**	**0.345**
Combination Median	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.052**	**0.345**
Combination Trim1	0.007	**0.895**	0.027	**0.539**	0.023	0.024	0.025	**0.959**	**0.943**	**0.929**	**0.074**	**0.157**	**0.268**	0.001	**0.050**
Combination Trim2	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.052**	**0.345**

Notes: See Table B1.

Table B4. Backtesting results BCTI

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.432**	**0.395**	**0.511**	**0.665**	**0.429**	**0.443**	**0.426**	**0.464**	**0.952**	**0.394**	**0.567**	**0.676**	**0.483**	**0.325**	**0.234**
Historical Simulation-250	**0.431**	0.000	0.000	**0.606**	**0.563**	**0.603**	**0.592**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation-500	**0.932**	0.000	0.000	**0.848**	**0.915**	**0.907**	**0.904**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical SimulationAll	0.013	0.000	0.000	**0.060**	**0.097**	**0.150**	**0.163**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.000	0.027	0.000	0.000	0.001	0.002	0.002	0.033	0.023	**0.141**	0.000	0.000	0.001	0.000	**0.081**
Variance-Covariance	0.000	0.037	0.000	0.000	0.000	0.000	0.001	**0.066**	0.001	0.004	0.000	0.000	0.000	0.000	0.001
Extreme Value Theory	0.000	**0.772**	0.000	0.000	0.000	0.000	0.000	**0.701**	**0.570**	**0.303**	0.000	0.000	0.000	0.000	**0.603**
AR-ARCH(1)-N	0.000	**0.534**	0.000	0.000	0.000	0.000	0.000	**0.814**	**0.587**	**0.738**	0.000	0.000	0.000	0.000	**0.051**
AR-ARCH(1)-t	0.008	**0.902**	0.030	0.041	0.013	0.024	0.020	**0.921**	**0.119**	**0.411**	0.045	0.009	0.016	0.019	**0.522**
AR-GARCH(1,1)-N	0.000	**0.564**	0.000	0.000	0.000	0.000	0.000	**0.833**	**0.764**	**0.711**	0.000	0.000	0.000	0.000	**0.091**
AR-GARCH(1,1)-t	0.020	**0.803**	**0.065**	**0.070**	0.027	0.027	0.022	**0.839**	**0.772**	**0.418**	**0.086**	**0.178**	**0.210**	0.040	**0.472**
AR-EGARCH(1,1)-N	0.000	**0.423**	0.000	0.000	0.000	0.000	0.000	**0.731**	**0.774**	**0.955**	0.000	0.000	0.000	0.000	**0.287**
AR-EGARCH(1,1)-t	0.020	**0.803**	**0.065**	**0.086**	0.027	0.031	0.027	**0.839**	**0.792**	**0.733**	**0.086**	**0.158**	**0.190**	0.048	**0.680**
Monte Carlo Simulation	0.030	**0.756**	**0.092**	**0.120**	0.038	0.043	0.037	**0.799**	**0.855**	**0.665**	**0.116**	**0.208**	**0.235**	**0.069**	**0.657**
Combination Mean	0.000	**0.423**	0.000	0.000	0.000	0.000	0.000	**0.731**	**0.287**	**0.460**	0.000	0.000	0.000	0.000	**0.074**
Combination Median	0.000	**0.491**	0.000	0.000	0.000	0.000	0.000	**0.601**	**0.465**	**0.744**	0.000	0.000	0.001	0.000	**0.465**
Combination Trim1	0.000	**0.349**	0.000	0.000	0.000	0.000	0.000	**0.664**	**0.409**	**0.630**	0.000	0.000	0.000	0.000	**0.171**
Combination Trim2	0.000	**0.491**	0.000	0.000	0.000	0.000	0.000	**0.601**	**0.155**	**0.370**	0.000	0.000	0.000	0.000	**0.071**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.930**	**0.658**	**0.903**	**0.126**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.591**	**0.347**
Historical Simulation-250	**0.054**	0.029	0.014	**0.090**	**0.059**	**0.110**	**0.093**	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003
Historical Simulation-500	**0.269**	0.012	0.023	**0.377**	**0.311**	**0.331**	**0.352**	0.000	0.000	0.000	0.000	0.000	0.000	0.031	0.012
Historical SimulationAll	**0.682**	**0.064**	**0.165**	**0.274**	**0.719**	**0.746**	**0.741**	0.002	0.000	0.002	0.008	0.001	0.001	**0.230**	**0.122**
Riskmetrics	**0.682**	**0.690**	**0.849**	**0.655**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.622**	**0.526**
Variance-Covariance	**0.146**	0.024	0.027	0.026	**0.261**	**0.335**	**0.311**	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.048
Extreme Value Theory	**0.269**	**0.534**	**0.447**	**0.207**	**0.237**	**0.230**	**0.222**	**0.606**	**0.462**	**0.364**	**0.444**	**0.592**	**0.699**	**0.154**	**0.079**
AR-ARCH(1)-N	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.062**	**0.455**
AR-ARCH(1)-t	**0.146**	**0.791**	**0.335**	0.026	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.175**	**0.717**
AR-GARCH(1,1)-N	**0.275**	**0.757**	**0.526**	**0.075**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.348**	**0.906**
AR-GARCH(1,1)-t	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**
AR-EGARCH(1,1)-N	**0.457**	**0.723**	**0.712**	**0.161**	**0.471**	**0.470**	**0.468**	**0.824**	**0.752**	**0.698**	**0.756**	**0.895**	**0.957**	**0.339**	**0.322**
AR-EGARCH(1,1)-t	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**
Monte Carlo Simulation	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**
Combination Mean	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**
Combination Median	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**
Combination Trim1	0.025	**0.860**	**0.079**	**0.231**	0.049	**0.050**	**0.050**	**0.938**	**0.911**	**0.891**	**0.144**	**0.275**	**0.423**	0.017	**0.313**
Combination Trim2	**0.066**	**0.825**	**0.180**	0.006	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.057**	**0.392**

Notes: See Table B1.

Table B5. Backtesting results 4 Time Charter Average Capesize

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	0.023	**0.238**	0.038	0.002	0.023	0.028	0.026	**0.150**	0.046	**0.152**	0.021	0.022	0.043	**0.111**	**0.788**
Historical Simulation -One Year Data	**0.061**	0.000	0.000	0.006	**0.251**	**0.275**	**0.288**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation -Two Year Data	0.000	0.000	0.000	0.000	0.041	**0.058**	**0.068**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation All Sample	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.623**	0.000	0.000	**0.102**	**0.750**	**0.785**	**0.785**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.000	0.000	0.000	0.000	**0.829**
AR-ARCH1-Normal Distribution	0.000	**0.666**	0.000	0.000	0.000	0.000	0.000	**0.593**	**0.068**	**0.121**	0.000	0.000	0.000	0.000	**0.661**
AR-ARCH1-Student Distribution	0.000	**0.486**	0.000	0.000	0.000	0.000	0.000	**0.355**	**0.463**	**0.697**	0.000	0.000	0.000	0.000	**0.756**
AR-GARCH11-Normal Distibution	**0.820**	**0.190**	**0.412**	0.031	**0.827**	**0.829**	**0.828**	**0.142**	0.043	0.038	**0.332**	**0.244**	**0.326**	**0.727**	**0.427**
AR-GARCH11-Student Distribution	**0.230**	**0.076**	**0.100**	0.013	**0.249**	**0.248**	**0.249**	0.033	**0.095**	**0.126**	0.048	**0.106**	**0.185**	**0.551**	**0.731**
AR-EGARCH11-Normal Distibution	**0.610**	**0.244**	**0.446**	0.019	**0.621**	**0.629**	**0.637**	**0.191**	**0.070**	0.030	**0.372**	**0.313**	**0.286**	**0.423**	**0.207**
AR-EGARCH11-Student Distribution	0.023	**0.238**	0.038	0.000	0.023	0.026	0.026	**0.150**	**0.111**	**0.150**	0.021	0.038	**0.076**	**0.102**	**0.617**
Montecarlo Simulation	**0.230**	**0.193**	**0.208**	0.013	**0.240**	**0.240**	**0.241**	**0.127**	**0.207**	**0.232**	**0.147**	**0.272**	**0.406**	**0.570**	**0.817**
Combination Mean	**0.516**	0.001	0.002	**0.655**	**0.567**	**0.578**	**0.569**	0.000	0.000	0.001	0.000	0.000	0.000	**0.193**	**0.081**
Combination Median	**0.610**	**0.244**	**0.446**	**0.250**	**0.621**	**0.622**	**0.624**	**0.191**	**0.170**	**0.139**	**0.372**	**0.500**	**0.601**	**0.847**	**0.652**
Combination Trim1	**0.516**	0.003	0.009	**0.525**	**0.559**	**0.564**	**0.557**	0.000	0.001	0.008	0.001	0.002	0.004	**0.242**	**0.107**
Combination Trim2	**0.431**	0.004	0.010	**0.370**	**0.475**	**0.481**	**0.473**	0.000	0.001	0.012	0.001	0.002	0.006	**0.198**	**0.092**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	0.029	**0.319**	**0.056**	**0.069**	0.020	0.025	0.031	**0.102**	0.023	0.006	0.014	0.012	0.010	0.024	**0.052**
Historical Simulation -One Year Data	**0.054**	0.002	0.001	0.007	**0.078**	**0.129**	**0.153**	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001
Historical Simulation -Two Year Data	0.003	0.000	0.000	0.002	0.034	**0.057**	**0.061**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation All Sample	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.029	0.003	0.001	0.015	0.040	**0.058**	**0.061**	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	**0.320**	0.000	0.000	0.000	0.000	0.000	**0.051**	0.007	0.000	0.000	0.000	0.000	0.000	**0.271**
AR-ARCH1-Normal Distribution	0.000	**0.279**	0.000	0.000	0.000	0.000	0.000	**0.037**	**0.236**	**0.460**	0.000	0.000	0.000	0.000	**0.336**
AR-ARCH1-Student Distribution	0.000	**0.152**	0.000	0.000	0.000	0.000	0.000	**0.072**	**0.994**	**0.270**	0.000	0.000	0.000	0.000	**0.841**
AR-GARCH11-Normal Distibution	0.000	**0.738**	0.000	0.000	0.000	0.000	0.000	**0.569**	0.000	0.000	0.000	0.000	0.000	0.000	**0.133**
AR-GARCH11-Student Distribution	**0.097**	**0.477**	**0.196**	**0.214**	**0.069**	**0.083**	**0.078**	**0.528**	**0.365**	**0.671**	**0.164**	**0.066**	**0.110**	**0.220**	**0.288**
AR-EGARCH11-Normal Distibution	0.000	**0.789**	0.000	0.000	0.000	0.000	0.000	**0.648**	0.001	0.000	0.000	0.000	0.000	0.000	**0.228**
AR-EGARCH11-Student Distribution	0.007	**0.373**	0.018	0.021	0.002	0.003	0.002	**0.377**	**0.720**	**0.857**	0.007	0.008	0.013	0.050	**0.766**
Montecarlo Simulation	**0.097**	**0.477**	**0.196**	**0.214**	**0.069**	**0.083**	**0.078**	**0.528**	**0.365**	**0.671**	**0.164**	**0.066**	**0.110**	**0.226**	**0.298**
Combination Mean	**0.413**	**0.564**	**0.605**	**0.641**	**0.388**	**0.458**	**0.451**	**0.645**	0.001	0.008	**0.625**	0.000	0.000	**0.070**	0.025
Combination Median	0.015	**0.398**	0.035	0.021	0.005	0.017	0.016	**0.414**	0.000	0.003	0.017	0.000	0.000	0.035	**0.161**
Combination Trim1	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.719**	**0.719**	**0.791**	0.044	**0.114**	**0.891**	0.020	0.043	**0.259**	**0.143**
Combination Trim2	**0.054**	**0.450**	**0.118**	**0.126**	0.032	**0.055**	**0.051**	**0.490**	0.012	**0.070**	**0.086**	0.000	0.000	**0.084**	**0.135**

Notes: See Table B1.

Table B6. Backtesting results 4 Time Charter Average Panamax

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.820**	**0.190**	**0.412**	**0.868**	**0.827**	**0.829**	**0.824**	**0.142**	0.043	**0.087**	**0.332**	**0.244**	**0.384**	**0.701**	**0.401**
Historical Simulation -One Year Data	**0.061**	0.000	0.000	0.000	**0.566**	**0.535**	**0.603**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation -Two Year Data	0.002	0.000	0.000	0.000	**0.226**	**0.249**	**0.257**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation All Sample	0.000	0.000	0.000	0.000	0.004	0.005	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.712**	0.000	0.000	0.036	**0.839**	**0.838**	**0.830**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.051	0.041	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.378**
AR-ARCH1-Normal Distribution	**0.729**	**0.295**	**0.544**	**0.378**	**0.741**	**0.749**	**0.767**	**0.264**	**0.278**	**0.059**	**0.505**	**0.677**	**0.330**	**0.174**	**0.071**
AR-ARCH1-Student Distribution	0.033	**0.849**	**0.100**	0.024	0.025	0.031	0.042	**0.832**	**0.353**	**0.050**	**0.081**	**0.061**	0.018	**0.089**	**0.309**
AR-GARCH11-Normal Distibution	**0.278**	**0.431**	**0.407**	**0.574**	**0.303**	**0.307**	**0.323**	**0.430**	**0.544**	**0.345**	**0.418**	**0.630**	**0.675**	**0.458**	**0.744**
AR-GARCH11-Student Distribution	**0.954**	**0.366**	**0.664**	**0.853**	**0.956**	**0.963**	**0.969**	**0.333**	**0.071**	**0.105**	**0.624**	**0.322**	**0.474**	**0.813**	**0.541**
AR-EGARCH11-Normal Distibution	**0.729**	0.002	0.007	**0.970**	**0.763**	**0.767**	**0.775**	0.000	0.005	0.010	0.001	0.002	0.005	**0.443**	**0.237**
AR-EGARCH11-Student Distribution	**0.841**	0.002	0.009	**0.813**	**0.860**	**0.864**	**0.871**	0.000	0.006	0.014	0.001	0.003	0.008	**0.326**	**0.145**
Montecarlo Simulation	**0.954**	**0.144**	**0.343**	**0.474**	**0.957**	**0.963**	**0.969**	**0.104**	**0.074**	**0.109**	**0.265**	**0.330**	**0.483**	**0.763**	**0.478**
Combination Mean	0.008	0.001	0.000	0.017	0.035	**0.058**	**0.098**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Combination Median	0.045	0.012	0.006	**0.147**	**0.085**	**0.088**	**0.104**	0.006	0.032	0.015	0.004	0.011	0.015	0.046	**0.192**
Combination Trim1	0.013	0.001	0.000	0.022	0.045	**0.071**	**0.113**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Combination Trim2	0.013	0.001	0.000	0.010	0.045	**0.062**	**0.093**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.269**	**0.534**	**0.447**	**0.485**	**0.237**	**0.230**	**0.257**	**0.606**	**0.462**	**0.499**	**0.444**	**0.592**	**0.156**	**0.559**	**0.479**
Historical Simulation -One Year Data	0.029	0.000	0.000	0.015	**0.172**	**0.253**	**0.201**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation -Two Year Data	**0.054**	0.000	0.000	0.007	**0.314**	**0.282**	**0.333**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation All Sample	0.001	0.000	0.000	0.000	**0.216**	**0.173**	**0.206**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.015	0.000	0.000	0.036	**0.108**	**0.119**	**0.143**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variance-Covariance	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.007	**0.373**	0.018	0.021	0.002	0.003	0.004	**0.377**	**0.720**	**0.266**	0.007	0.008	0.007	**0.052**	**0.963**
AR-ARCH1-Normal Distribution	0.003	**0.349**	0.008	0.002	0.001	0.001	0.001	**0.342**	**0.823**	**0.350**	0.002	0.003	0.004	0.005	**0.072**
AR-ARCH1-Student Distribution	**0.166**	**0.505**	**0.307**	0.026	**0.134**	**0.127**	**0.121**	**0.567**	**0.414**	**0.313**	**0.284**	**0.413**	**0.520**	0.013	0.006
AR-GARCH11-Normal Distibution	**0.054**	**0.450**	**0.118**	**0.125**	0.032	0.040	**0.050**	**0.490**	**0.443**	**0.097**	**0.086**	0.046	0.023	**0.256**	**0.929**
AR-GARCH11-Student Distribution	**0.823**	**0.626**	**0.866**	**0.555**	**0.820**	**0.815**	**0.810**	**0.720**	**0.610**	**0.529**	**0.914**	**0.958**	**0.978**	**0.804**	**0.513**
AR-EGARCH11-Normal Distibution	**0.097**	**0.477**	**0.196**	**0.130**	**0.069**	**0.083**	**0.099**	**0.528**	**0.365**	**0.064**	**0.164**	**0.066**	**0.026**	**0.194**	**0.240**
AR-EGARCH11-Student Distribution	**0.930**	**0.658**	**0.903**	**0.788**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.707**	**0.468**
Montecarlo Simulation	**0.930**	**0.658**	**0.903**	**0.664**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.332**	**0.153**
Combination Mean	**0.457**	0.048	**0.107**	**0.679**	**0.529**	**0.575**	**0.613**	0.001	0.000	0.000	0.003	0.000	0.000	**0.081**	**0.050**
Combination Median	**0.275**	**0.757**	**0.526**	**0.332**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.257**	**0.432**
Combination Trim1	0.025	0.009	0.003	**0.108**	**0.140**	**0.239**	**0.327**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Combination Trim2	**0.682**	**0.064**	**0.165**	**0.714**	**0.719**	**0.746**	**0.769**	0.002	0.000	0.000	0.008	0.001	0.000	**0.089**	0.038

Notes: See Table B1.

Table B7. Backtesting results TD3

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.278**	0.002	0.005	**0.317**	**0.353**	**0.361**	**0.399**	0.000	0.003	0.000	0.001	0.003	0.001	**0.325**	**0.373**
Historical Simulation -One Year Data	**0.217**	0.000	0.000	**0.293**	**0.398**	**0.421**	**0.477**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation -Two Year Data	**0.064**	0.000	0.000	**0.202**	**0.203**	**0.255**	**0.269**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Historical Simulation All Sample	0.045	0.000	0.000	**0.158**	**0.191**	**0.206**	**0.235**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	**0.432**	0.000	0.000	**0.771**	**0.574**	**0.593**	**0.598**	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.004
Variance-Covariance	0.000	0.000	0.000	0.000	0.003	0.007	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Extreme Value Theory	0.000	**0.053**	0.000	0.000	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	**0.249**
AR-ARCH1-Normal Distribution	0.000	**0.373**	0.000	0.000	0.000	0.000	0.000	**0.687**	**0.366**	**0.255**	0.000	0.000	0.000	0.000	**0.410**
AR-ARCH1-Student Distribution	**0.523**	**0.363**	**0.539**	**0.345**	**0.519**	**0.559**	**0.581**	**0.431**	**0.185**	**0.061**	**0.602**	0.049	0.044	**0.717**	**0.781**
AR-GARCH11-Normal Distibution	0.001	**0.063**	0.001	0.006	0.005	0.009	0.020	**0.071**	0.014	0.001	0.002	0.001	0.000	0.003	**0.935**
AR-GARCH11-Student Distribution	**0.841**	**0.765**	**0.937**	**0.915**	**0.840**	**0.857**	**0.867**	**0.774**	**0.139**	**0.067**	**0.940**	**0.125**	**0.141**	**0.810**	**0.579**
AR-EGARCH11-Normal Distibution	0.000	0.027	0.000	0.000	0.001	0.002	0.005	0.033	0.023	0.003	0.000	0.000	0.000	0.000	**0.557**
AR-EGARCH11-Student Distribution	**0.623**	**0.612**	**0.779**	**0.050**	**0.632**	**0.650**	**0.646**	**0.607**	**0.232**	**0.621**	**0.778**	**0.595**	**0.616**	**0.739**	**0.641**
Montecarlo Simulation	**0.350**	**0.982**	**0.646**	**0.645**	**0.360**	**0.399**	**0.400**	**0.983**	**0.131**	**0.098**	**0.657**	**0.150**	**0.214**	**0.496**	**0.596**
Combination Mean	0.020	0.001	0.000	0.029	**0.058**	**0.068**	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.030	**0.288**
Combination Median	0.003	**0.090**	0.003	0.008	0.010	0.016	0.029	**0.099**	0.024	0.002	0.005	0.005	0.001	0.009	**0.880**
Combination Trim1	0.020	0.001	0.000	0.029	**0.058**	**0.068**	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.030	**0.288**
Combination Trim2	0.020	0.001	0.000	0.029	**0.058**	**0.068**	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.030	**0.288**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.066**	**0.825**	**0.180**	**0.043**	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.053**	**0.355**
Historical Simulation -One Year Data	**0.413**	0.009	0.023	**0.641**	**0.465**	**0.448**	**0.444**	0.000	0.002	0.015	0.000	0.000	0.000	**0.298**	**0.143**
Historical Simulation -Two Year Data	**0.823**	0.004	0.016	**0.929**	**0.852**	**0.844**	**0.841**	0.000	0.000	0.004	0.000	0.000	0.000	**0.449**	**0.207**
Historical Simulation All Sample	**0.066**	0.015	0.010	**0.152**	**0.185**	**0.167**	**0.171**	0.000	0.006	0.018	0.000	0.000	0.000	**0.061**	**0.447**
Riskmetrics	0.001	0.000	0.000	0.005	0.004	0.005	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.005	**0.172**
Variance-Covariance	**0.682**	0.002	0.006	**0.951**	**0.753**	**0.741**	**0.744**	0.000	0.000	0.001	0.000	0.000	0.000	**0.141**	**0.067**
Extreme Value Theory	**0.413**	**0.564**	**0.605**	**0.452**	**0.388**	**0.380**	**0.410**	**0.645**	**0.511**	**0.407**	**0.625**	**0.761**	**0.158**	**0.743**	**0.572**
AR-ARCH1-Normal Distribution	**0.823**	**0.626**	**0.866**	**0.827**	**0.820**	**0.815**	**0.827**	**0.720**	**0.610**	**0.250**	**0.914**	**0.958**	**0.112**	**0.894**	**0.643**
AR-ARCH1-Student Distribution	**0.146**	**0.791**	**0.335**	**0.328**	**0.178**	**0.177**	**0.176**	**0.885**	**0.838**	**0.801**	**0.402**	**0.607**	**0.763**	**0.176**	**0.732**
AR-GARCH11-Normal Distibution	**0.823**	**0.103**	**0.258**	**0.827**	**0.836**	**0.846**	**0.854**	0.007	0.000	0.000	0.026	0.004	0.001	**0.761**	**0.463**
AR-GARCH11-Student Distribution	**0.457**	**0.723**	**0.712**	**0.464**	**0.471**	**0.470**	**0.526**	**0.824**	**0.752**	**0.092**	**0.756**	**0.895**	0.019	**0.395**	**0.410**
AR-EGARCH11-Normal Distibution	**0.823**	**0.103**	**0.258**	**0.827**	**0.836**	**0.846**	**0.869**	0.007	0.000	0.000	0.026	0.004	0.000	**0.245**	**0.094**
AR-EGARCH11-Student Distribution	**0.146**	**0.791**	**0.335**	**0.127**	**0.178**	**0.177**	**0.259**	**0.885**	**0.838**	0.032	**0.402**	**0.607**	0.001	**0.136**	**0.422**
Montecarlo Simulation	**0.275**	**0.757**	**0.526**	**0.075**	**0.301**	**0.299**	**0.372**	**0.856**	**0.796**	**0.058**	**0.580**	**0.773**	0.006	**0.313**	**0.634**
Combination Mean	0.025	**0.860**	**0.079**	0.010	0.049	**0.050**	**0.050**	**0.938**	**0.911**	**0.891**	**0.144**	**0.275**	**0.423**	0.023	**0.555**
Combination Median	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.076**	**0.715**
Combination Trim1	0.025	**0.860**	**0.079**	0.010	0.049	**0.050**	**0.050**	**0.938**	**0.911**	**0.891**	**0.144**	**0.275**	**0.423**	0.023	**0.555**
Combination Trim2	**0.066**	**0.825**	**0.180**	0.043	**0.097**	**0.097**	**0.097**	**0.913**	**0.876**	**0.848**	**0.253**	**0.431**	**0.599**	**0.076**	**0.715**

Notes: See Table B1.

Table B8. Backtesting results TD5

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.431**	**0.138**	**0.245**	**0.465**	**0.405**	**0.390**	**0.392**	**0.177**	**0.133**	**0.349**	**0.291**	**0.377**	**0.504**	**0.052**	0.019
Historical Simulation -One Year Data	**0.523**	**0.232**	**0.400**	**0.855**	**0.546**	**0.576**	**0.577**	**0.205**	0.026	**0.084**	**0.367**	**0.132**	**0.226**	**0.735**	**0.871**
Historical Simulation -Two Year Data	**0.623**	**0.612**	**0.779**	**0.548**	**0.632**	**0.665**	**0.667**	**0.607**	0.032	**0.095**	**0.778**	**0.074**	**0.136**	**0.722**	**0.607**
Historical Simulation All Sample	**0.141**	0.038	0.039	**0.209**	**0.160**	**0.181**	**0.185**	0.010	0.002	0.004	0.012	0.007	0.015	**0.409**	**0.755**
Riskmetrics	**0.064**	**0.672**	**0.165**	**0.160**	**0.081**	**0.115**	**0.103**	**0.697**	0.034	**0.250**	**0.192**	0.017	0.021	**0.116**	**0.494**
Variance-Covariance	**0.432**	**0.518**	**0.596**	**0.533**	**0.450**	**0.495**	**0.507**	**0.515**	0.018	0.023	**0.600**	0.039	**0.071**	**0.487**	**0.437**
Extreme Value Theory	0.000	**0.369**	0.000	0.000	0.000	0.000	0.000	**0.202**	0.015	0.008	0.000	0.000	0.000	0.000	**0.409**
AR-ARCH1-Normal Distribution	**0.932**	0.019	**0.062**	**0.651**	**0.928**	**0.925**	**0.922**	**0.086**	**0.900**	**0.622**	**0.228**	**0.166**	**0.185**	**0.646**	**0.351**
AR-ARCH1-Student Distribution	0.011	**0.129**	0.013	0.031	0.006	0.007	0.010	**0.116**	**0.824**	**0.550**	0.008	0.011	0.011	0.033	**0.241**
AR-GARCH11-Normal Distibution	**0.623**	0.030	**0.084**	**0.548**	**0.610**	**0.643**	**0.649**	**0.128**	**0.457**	**0.512**	**0.280**	0.034	0.070	**0.743**	**0.649**
AR-GARCH11-Student Distribution	0.016	**0.145**	0.019	0.021	0.009	0.009	0.012	**0.132**	**0.348**	**0.975**	0.014	0.034	0.032	0.027	**0.123**
AR-EGARCH11-Normal Distibution	**0.432**	0.038	**0.085**	**0.344**	**0.419**	**0.454**	**0.460**	**0.154**	**0.659**	**0.659**	**0.269**	**0.090**	**0.164**	**0.562**	**0.571**
AR-EGARCH11-Student Distribution	**0.108**	**0.271**	**0.150**	**0.159**	**0.087**	**0.076**	**0.086**	**0.267**	**0.110**	**0.568**	**0.136**	**0.149**	**0.146**	0.026	0.022
Montecarlo Simulation	0.016	**0.145**	0.019	0.021	0.009	0.008	0.011	**0.132**	**0.177**	**0.751**	0.014	0.032	0.031	0.016	**0.064**
Combination Mean	**0.932**	**0.231**	**0.486**	**0.848**	**0.929**	**0.929**	**0.924**	**0.287**	**0.415**	**0.376**	**0.565**	**0.120**	**0.201**	**0.270**	**0.106**
Combination Median	**0.623**	0.030	**0.084**	**0.754**	**0.610**	**0.628**	**0.632**	**0.128**	**0.811**	**0.818**	**0.280**	**0.284**	**0.435**	**0.092**	0.036
Combination Trim1	**0.932**	**0.231**	**0.486**	**0.848**	**0.929**	**0.929**	**0.924**	**0.287**	**0.415**	**0.376**	**0.565**	**0.120**	**0.201**	**0.405**	**0.179**
Combination Trim2	**0.820**	**0.209**	**0.443**	**0.787**	**0.813**	**0.819**	**0.815**	**0.262**	**0.476**	**0.446**	**0.520**	**0.131**	**0.221**	**0.210**	**0.078**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.823**	**0.626**	**0.866**	**0.929**	**0.820**	**0.815**	**0.810**	**0.720**	**0.610**	**0.529**	**0.914**	**0.958**	**0.978**	**0.403**	**0.178**
Historical Simulation -One Year Data	**0.599**	**0.126**	**0.270**	**0.421**	**0.617**	**0.670**	**0.685**	0.012	0.000	0.000	0.037	0.000	0.000	**0.519**	**0.261**
Historical Simulation -Two Year Data	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.719**	**0.719**	**0.791**	0.044	**0.114**	**0.891**	0.020	0.043	**0.432**	**0.288**
Historical Simulation All Sample	**0.823**	**0.626**	**0.866**	**0.929**	**0.820**	**0.832**	**0.828**	**0.720**	**0.098**	**0.223**	**0.914**	**0.061**	**0.113**	**0.775**	**0.479**
Riskmetrics	0.029	**0.423**	**0.066**	0.044	0.014	0.018	0.016	**0.451**	**0.529**	**0.902**	0.040	0.028	0.048	0.018	0.036
Variance-Covariance	0.001	**0.086**	0.001	0.003	0.001	0.004	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.010
Extreme Value Theory	**0.166**	**0.505**	**0.307**	**0.259**	**0.134**	**0.127**	**0.121**	**0.567**	**0.414**	**0.313**	**0.284**	**0.413**	**0.520**	**0.541**	**0.888**
AR-ARCH1-Normal Distribution	0.015	**0.398**	0.035	0.036	0.005	0.011	0.015	**0.414**	0.031	0.007	0.017	0.000	0.000	**0.093**	**0.881**
AR-ARCH1-Student Distribution	**0.413**	**0.564**	**0.605**	**0.641**	**0.388**	**0.380**	**0.410**	**0.645**	**0.511**	**0.407**	**0.625**	**0.761**	**0.158**	**0.847**	**0.810**
AR-GARCH11-Normal Distibution	**0.097**	**0.477**	**0.196**	**0.153**	**0.069**	**0.105**	**0.100**	**0.528**	0.007	0.044	**0.164**	0.000	0.000	**0.347**	**0.638**
AR-GARCH11-Student Distribution	**0.823**	**0.626**	**0.866**	**0.929**	**0.820**	**0.815**	**0.810**	**0.720**	**0.610**	**0.529**	**0.914**	**0.958**	**0.978**	**0.572**	**0.292**
AR-EGARCH11-Normal Distibution	**0.269**	**0.534**	**0.447**	**0.406**	**0.237**	**0.265**	**0.257**	**0.606**	**0.234**	**0.464**	**0.444**	**0.094**	**0.158**	**0.708**	**0.867**
AR-EGARCH11-Student Distribution	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.715**	**0.727**
Montecarlo Simulation	**0.413**	**0.564**	**0.605**	**0.587**	**0.388**	**0.380**	**0.372**	**0.645**	**0.511**	**0.417**	**0.625**	**0.761**	**0.845**	**0.700**	**0.507**
Combination Mean	**0.823**	**0.626**	**0.866**	**0.929**	**0.820**	**0.815**	**0.810**	**0.720**	**0.610**	**0.529**	**0.914**	**0.958**	**0.978**	**0.996**	**0.981**
Combination Median	**0.930**	**0.658**	**0.903**	**0.998**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.894**	**0.810**
Combination Trim1	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.715**	**0.727**
Combination Trim2	**0.599**	**0.594**	**0.756**	**0.775**	**0.585**	**0.579**	**0.572**	**0.683**	**0.560**	**0.473**	**0.795**	**0.889**	**0.937**	**0.741**	**0.459**

Notes: See Table B1.

Table B9 Backtesting results TD7

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.712**	**0.919**	**0.929**	**0.544**	**0.711**	**0.701**	**0.811**	**0.917**	**0.759**	**0.809**	**0.928**	**0.931**	**0.918**	**0.874**	**0.638**
Historical Simulation -One Year Data	**0.623**	0.000	0.000	**0.323**	**0.694**	**0.732**	**0.751**	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
Historical Simulation -Two Year Data	0.030	0.000	0.000	0.028	**0.084**	**0.125**	**0.156**	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004
Historical Simulation All Sample	**0.932**	0.000	0.000	**0.510**	**0.950**	**0.954**	**0.933**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Riskmetrics	0.001	**0.285**	0.002	0.005	0.003	0.004	0.000	**0.339**	**0.462**	**0.104**	0.006	0.016	0.009	0.003	**0.737**
Variance-Covariance	0.030	0.000	0.000	0.028	**0.084**	**0.138**	**0.179**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.090
AR-ARCH1-Normal Distribution	0.003	**0.152**	0.004	0.006	0.004	0.004	0.004	**0.412**	**0.238**	**0.961**	0.015	0.028	0.015	0.009	**0.979**
AR-ARCH1-Student Distribution	**0.354**	**0.425**	**0.474**	**0.350**	**0.334**	**0.355**	**0.440**	**0.430**	**0.520**	**0.376**	**0.469**	**0.228**	**0.347**	**0.720**	**0.792**
AR-GARCH11-Normal Distibution	0.001	**0.179**	0.002	0.006	0.002	0.003	0.000	**0.457**	**0.880**	**0.220**	0.007	0.013	0.004	0.003	**0.743**
AR-GARCH11-Student Distribution	**0.354**	**0.872**	**0.643**	**0.156**	**0.343**	**0.337**	**0.418**	**0.866**	**0.809**	**0.904**	**0.631**	**0.810**	**0.919**	**0.559**	**0.449**
AR-EGARCH11-Normal Distibution	0.001	**0.194**	0.001	0.003	0.001	0.002	0.003	**0.480**	**0.822**	**0.184**	0.005	0.008	0.002	0.002	**0.965**
AR-EGARCH11-Student Distribution	**0.610**	**0.244**	**0.446**	**0.385**	**0.621**	**0.629**	**0.744**	**0.191**	**0.070**	0.026	**0.372**	**0.313**	**0.274**	**0.935**	**0.949**
Montecarlo Simulation	**0.610**	**0.972**	**0.878**	**0.149**	**0.607**	**0.595**	**0.694**	**0.971**	**0.689**	**0.804**	**0.875**	**0.888**	**0.975**	**0.716**	**0.463**
Combination Mean	**0.064**	**0.079**	0.039	**0.161**	**0.066**	**0.087**	**0.086**	**0.268**	**0.575**	**0.099**	**0.112**	**0.051**	0.014	**0.146**	**0.972**
Combination Median	0.013	**0.116**	0.013	**0.057**	0.016	0.013	0.016	**0.347**	**0.175**	**0.806**	0.041	**0.062**	0.013	0.031	**0.612**
Combination Trim1	**0.064**	**0.079**	0.039	**0.161**	**0.066**	**0.079**	**0.078**	**0.268**	**0.984**	**0.234**	**0.112**	**0.135**	0.039	**0.145**	**0.911**
Combination Trim2	**0.064**	**0.079**	0.039	**0.161**	**0.066**	**0.079**	**0.078**	**0.268**	**0.984**	**0.234**	**0.112**	**0.135**	0.039	**0.145**	**0.911**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.275**	**0.757**	**0.526**	**0.559**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.334**	**0.761**
Historical Simulation -One Year Data	**0.413**	**0.152**	**0.256**	**0.292**	**0.426**	**0.453**	**0.477**	0.019	0.001	0.000	0.045	0.011	0.004	**0.439**	**0.241**
Historical Simulation -Two Year Data	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.705**	**0.699**
Historical Simulation All Sample	**0.599**	**0.594**	**0.756**	**0.768**	**0.585**	**0.579**	**0.637**	**0.683**	**0.560**	0.007	**0.795**	**0.889**	0.000	**0.327**	**0.139**
Riskmetrics	**0.097**	**0.477**	**0.196**	**0.214**	**0.069**	**0.064**	**0.060**	**0.528**	**0.368**	**0.266**	**0.164**	**0.257**	**0.345**	0.046	0.039
Variance-Covariance	0.007	**0.058**	0.004	0.021	0.005	0.015	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Extreme Value Theory	**0.097**	**0.477**	**0.196**	**0.197**	**0.069**	**0.064**	**0.118**	**0.528**	**0.368**	**0.297**	**0.164**	**0.257**	**0.509**	**0.342**	**0.619**
AR-ARCH1-Normal Distribution	0.029	**0.423**	**0.066**	0.015	0.014	0.012	0.023	**0.451**	**0.282**	**0.139**	0.040	**0.071**	0.000	0.032	**0.074**
AR-ARCH1-Student Distribution	**0.269**	**0.534**	**0.447**	**0.025**	**0.237**	**0.230**	**0.257**	**0.606**	**0.462**	**0.499**	**0.444**	**0.592**	**0.156**	**0.487**	**0.379**
AR-GARCH11-Normal Distibution	**0.097**	**0.477**	**0.196**	**0.214**	**0.069**	**0.064**	**0.060**	**0.528**	**0.368**	**0.266**	**0.164**	**0.257**	**0.345**	**0.193**	**0.238**
AR-GARCH11-Student Distribution	**0.930**	**0.658**	**0.903**	**0.387**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.765**	**0.544**
AR-EGARCH11-Normal Distibution	**0.166**	**0.505**	**0.307**	**0.335**	**0.134**	**0.127**	**0.179**	**0.567**	**0.414**	0.040	**0.284**	**0.413**	0.000	**0.444**	**0.519**
AR-EGARCH11-Student Distribution	**0.413**	**0.564**	**0.605**	**0.641**	**0.388**	**0.380**	**0.410**	**0.645**	**0.511**	**0.407**	**0.625**	**0.761**	**0.158**	**0.603**	**0.390**
Montecarlo Simulation	**0.413**	**0.564**	**0.605**	**0.641**	**0.388**	**0.380**	**0.372**	**0.645**	**0.511**	**0.417**	**0.625**	**0.761**	**0.845**	**0.821**	**0.727**
Combination Mean	**0.413**	**0.564**	**0.605**	**0.587**	**0.388**	**0.380**	**0.450**	**0.645**	**0.511**	0.013	**0.625**	**0.761**	0.000	**0.568**	**0.354**
Combination Median	**0.599**	**0.594**	**0.756**	**0.775**	**0.585**	**0.579**	**0.604**	**0.683**	**0.560**	**0.324**	**0.795**	**0.889**	**0.142**	**0.898**	**0.683**
Combination Trim1	**0.682**	**0.690**	**0.849**	**0.951**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.630**	**0.539**
Combination Trim2	**0.413**	**0.564**	**0.605**	**0.587**	**0.388**	**0.380**	**0.450**	**0.645**	**0.511**	0.013	**0.625**	**0.761**	0.000	**0.568**	**0.354**

Notes: See Table B1.

Table B10. Backtesting results TD9

Panel A: 5%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.712**	**0.582**	**0.803**	**0.851**	**0.705**	**0.695**	**0.805**	**0.592**	**0.443**	**0.890**	**0.808**	**0.865**	**0.885**	**0.889**	**0.667**
Historical Simulation -One Year Data	**0.431**	0.000	0.000	**0.665**	**0.503**	**0.527**	**0.627**	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
Historical Simulation -Two Year Data	**0.954**	0.003	0.012	**0.700**	**0.959**	**0.965**	**0.972**	0.000	0.002	0.001	0.002	0.004	0.004	0.005	0.001
Historical Simulation All Sample	**0.431**	0.000	0.001	**0.192**	**0.494**	**0.495**	**0.520**	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
Riskmetrics	0.020	0.007	0.002	0.029	**0.050**	**0.054**	**0.052**	0.004	0.019	**0.074**	0.001	0.004	0.009	0.018	**0.139**
Variance-Covariance	**0.278**	0.000	0.001	**0.079**	**0.366**	**0.395**	**0.426**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Extreme Value Theory	0.000	**0.234**	0.000	0.000	0.000	0.000	0.000	**0.093**	**0.188**	**0.036**	0.000	0.000	0.000	0.000	0.700
AR-ARCH1-Normal Distribution	**0.123**	**0.773**	**0.293**	**0.067**	**0.141**	**0.157**	**0.163**	**0.790**	**0.374**	**0.383**	**0.319**	**0.354**	**0.510**	0.011	0.012
AR-ARCH1-Student Distribution	0.001	**0.542**	0.004	0.001	0.001	0.001	0.002	**0.450**	**0.773**	0.045	0.002	0.004	0.000	0.002	**0.072**
AR-GARCH11-Normal Distibution	**0.217**	**0.877**	**0.461**	0.027	**0.232**	**0.226**	**0.181**	**0.886**	**0.762**	**0.913**	**0.481**	**0.620**	**0.712**	**0.293**	**0.430**
AR-GARCH11-Student Distribution	0.002	**0.826**	0.006	0.004	0.001	0.001	0.002	**0.795**	**0.445**	**0.798**	0.003	0.007	0.010	0.013	**0.905**
AR-EGARCH11-Normal Distibution	**0.217**	**0.877**	**0.461**	**0.077**	**0.232**	**0.226**	**0.181**	**0.886**	**0.762**	**0.913**	**0.481**	**0.620**	**0.712**	**0.268**	**0.372**
AR-EGARCH11-Student Distribution	0.002	**0.826**	0.006	0.004	0.001	0.000	0.002	**0.795**	**0.249**	**0.784**	0.003	0.004	0.003	0.010	**0.521**
Montecarlo Simulation	0.001	**0.725**	0.003	0.002	0.000	0.000	0.001	**0.676**	**0.833**	**0.363**	0.001	0.003	0.003	0.006	**0.880**
Combination Mean	**0.841**	**0.765**	**0.937**	**0.299**	**0.840**	**0.847**	**0.741**	**0.774**	**0.950**	**0.954**	**0.940**	**0.967**	**0.982**	**0.766**	**0.518**
Combination Median	**0.523**	**0.914**	**0.811**	**0.345**	**0.528**	**0.540**	**0.454**	**0.919**	**0.735**	**0.785**	**0.816**	**0.865**	**0.916**	**0.385**	**0.251**
Combination Trim1	**0.954**	**0.762**	**0.954**	**0.272**	**0.955**	**0.961**	**0.853**	**0.757**	**0.665**	**0.561**	**0.952**	**0.979**	**0.985**	**0.662**	**0.375**
Combination Trim2	**0.712**	**0.490**	**0.736**	**0.179**	**0.716**	**0.712**	**0.817**	**0.455**	**0.553**	**0.550**	**0.706**	**0.872**	**0.952**	**0.598**	**0.322**

Panel B: 1%	LR_{uc}	LR_{ind}	LR_{cc}	LR_{muc}	$DQ1_{uc}$	$DQ2_{uc}$	$DQ3_{uc}$	$DQ1_{ind}$	$DQ2_{ind}$	$DQ3_{ind}$	$DQ1_{cc}$	$DQ2_{cc}$	$DQ3_{cc}$	Dur_{cc}	Dur_{ind}
Filtered Historical Simulation	**0.457**	**0.723**	**0.712**	**0.679**	**0.471**	**0.470**	**0.468**	**0.824**	**0.752**	**0.698**	**0.756**	**0.895**	**0.957**	**0.257**	**0.215**
Historical Simulation -One Year Data	**0.269**	**0.181**	**0.221**	**0.406**	**0.273**	**0.263**	**0.256**	0.029	**0.250**	**0.493**	0.046	**0.087**	**0.148**	**0.313**	**0.198**
Historical Simulation -Two Year Data	**0.413**	**0.152**	**0.256**	**0.587**	**0.426**	**0.453**	**0.477**	0.019	0.001	0.000	0.045	0.011	0.004	0.011	0.003
Historical Simulation All Sample	**0.166**	**0.212**	**0.176**	**0.259**	**0.161**	**0.216**	**0.235**	0.041	0.000	0.000	0.042	0.000	0.000	0.006	0.002
Riskmetrics	0.029	0.037	0.010	0.044	0.029	0.025	0.023	0.000	0.027	**0.136**	0.000	0.000	0.000	**0.102**	**0.342**
Variance-Covariance	0.003	**0.071**	0.003	0.008	0.002	0.004	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Extreme Value Theory	**0.097**	**0.477**	**0.196**	**0.061**	**0.069**	**0.064**	**0.060**	**0.528**	**0.368**	**0.266**	**0.164**	**0.257**	**0.345**	**0.341**	**0.613**
AR-ARCH1-Normal Distribution	0.007	**0.373**	0.018	0.010	0.002	0.002	0.001	**0.377**	**0.207**	**0.118**	0.007	0.012	0.019	**0.052**	**0.886**
AR-ARCH1-Student Distribution	**0.166**	**0.505**	**0.307**	**0.110**	**0.134**	**0.127**	**0.121**	**0.567**	**0.414**	**0.313**	**0.284**	**0.413**	**0.520**	**0.411**	**0.450**
AR-GARCH11-Normal Distibution	**0.054**	**0.450**	**0.118**	0.031	0.032	0.029	0.027	**0.490**	**0.324**	**0.222**	**0.086**	**0.143**	**0.202**	**0.253**	**0.861**
AR-GARCH11-Student Distribution	**0.097**	**0.477**	**0.196**	**0.061**	**0.069**	**0.064**	**0.060**	**0.528**	**0.368**	**0.266**	**0.164**	**0.257**	**0.345**	**0.341**	**0.613**
AR-EGARCH11-Normal Distibution	**0.054**	**0.450**	**0.118**	0.031	0.032	0.029	0.027	**0.490**	**0.324**	**0.222**	**0.086**	**0.143**	**0.202**	**0.257**	**0.992**
AR-EGARCH11-Student Distribution	**0.682**	**0.690**	**0.849**	**0.714**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.538**	**0.406**
Montecarlo Simulation	**0.269**	**0.534**	**0.447**	**0.186**	**0.237**	**0.230**	**0.222**	**0.606**	**0.462**	**0.364**	**0.444**	**0.592**	**0.699**	**0.718**	**0.984**
Combination Mean	**0.682**	**0.690**	**0.849**	**0.714**	**0.686**	**0.685**	**0.685**	**0.791**	**0.706**	**0.642**	**0.891**	**0.960**	**0.985**	**0.636**	**0.551**
Combination Median	**0.930**	**0.658**	**0.903**	**0.664**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.797**	**0.591**
Combination Trim1	**0.275**	**0.757**	**0.526**	**0.563**	**0.301**	**0.299**	**0.297**	**0.856**	**0.796**	**0.751**	**0.580**	**0.773**	**0.887**	**0.348**	**0.906**
Combination Trim2	**0.930**	**0.658**	**0.903**	**0.664**	**0.929**	**0.931**	**0.933**	**0.756**	**0.659**	**0.586**	**0.949**	**0.977**	**0.990**	**0.797**	**0.591**

Notes: See Table B1.

Table B11. Performance Evaluation: All methods

Panel A: 5%	4 TC CAPE			4 TC PAN			TD3			TD5			TD7			TD9		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.046	3.973	NR	0.014	3.381	NR	0.033	4.762	NR	0.053	4.528	NR	0.056	5.529	R	0.059	4.826	NR
Historical Simulation-250	0.054	4.589	NR	0.031	7.583	R	0.040	5.725	R	0.062	5.268	NR	0.062	6.150	NR	0.072	5.970	R
Historical Simulation-500	0.068	5.853	R	0.048	11.88	R	0.046	6.551	R	0.064	5.444	NR	0.056	5.542	NR	0.078	6.416	R
Historical SimulationAll	0.183	15.66	R	0.056	13.82	R	0.045	6.399	R	0.075	6.444	R	0.080	7.925	R	0.080	6.563	R
Riskmetrics	0.048	4.104	NR	0.021	5.241	NR	0.038	5.429	NR	0.052	4.461	NR	0.038	3.739	NR	0.044	3.588	NR
Variance-Covariance	0.162	13.92	R	0.048	11.86	R	0.046	6.611	R	0.066	5.631	NR	0.066	6.562	R	0.075	6.181	R
Extreme Value Theory	0.069	5.907	R	0.018	4.478	NR	0.064	9.047	R	0.108	9.237	R	0.108	10.720	R	0.111	9.121	R
AR-ARCH(1)-N	0.071	6.044	R	0.018	4.548	NR	0.036	5.081	NR	0.063	5.391	NR	0.046	4.580	NR	0.057	4.686	NR
AR-ARCH(1)-t	0.08	6.87	R	0.023	5.618	R	0.039	5.566	R	0.083	7.073	R	0.063	6.251	R	0.084	6.885	R
AR-GARCH(1.1)-N	0.043	3.649	NR	0.013	3.133	NR	0.035	4.998	NR	0.060	5.134	NR	0.039	3.910	NR	0.051	4.226	NR
AR-GARCH(1.1)-t	0.044	3.731	NR	0.013	3.312	NR	0.039	5.590	R	0.069	5.917	R	0.057	5.692	R	0.077	6.373	R
AR-EGARCH(1.1)-N	0.042	3.616	NR	0.012	3.058	NR	0.034	4.871	NR	0.056	4.816	NR	0.040	4.011	NR	0.050	4.107	NR
AR-EGARCH(1.1)-t	0.045	3.887	NR	0.012	3.074	NR	0.035	4.919	NR	0.055	4.734	NR	0.056	5.585	R	0.074	6.119	R
Monte Carlo Simulation	0.044	3.778	NR	0.013	3.277	NR	0.037	5.245	NR	0.066	5.664	R	0.055	5.463	R	0.080	6.555	R
Combination Mean	0.041	3.553	NR	0.018	4.407	NR	0.034	4.801	NR	0.059	5.067	NR	0.046	4.558	NR	0.055	4.549	NR
Combination Median	0.043	3.678	NR	0.012	3.063	NR	0.034	4.863	NR	0.056	4.819	NR	0.044	4.321	NR	0.054	4.485	NR
Combination Trim1	0.042	3.564	NR	0.017	4.246	NR	0.034	4.784	NR	0.061	5.177	NR	0.048	4.731	NR	0.056	4.599	NR
Combination Trim2	0.042	3.622	NR	0.016	4.009	NR	0.033	4.759	NR	0.061	5.197	NR	0.048	4.732	NR	0.058	4.753	NR
Panel B: 1%	4 TC CAPE			4 TC PAN			TD3			TD5			TD7			TD9		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.007	**2.580**	NR	0.002	1.989	NR	0.004	3.949	NR	0.006	3.254	NR	0.005	3.400	NR	0.006	3.561	NR
Historical Simulation-250	0.015	5.759	NR	0.005	5.631	NR	0.006	5.398	NR	0.012	7.028	NR	0.010	6.592	R	0.010	5.380	NR
Historical Simulation-500	0.012	4.546	NR	0.007	7.843	R	0.006	5.262	NR	0.012	7.047	NR	0.006	4.041	NR	0.017	9.689	R
Historical SimulationAll	0.040	15.973	R	0.012	13.714	R	0.005	4.915	NR	0.012	6.662	NR	0.008	5.335	NR	0.018	10.253	R
Riskmetrics	0.014	5.578	NR	0.004	4.078	NR	0.007	5.976	R	0.010	5.440	R	0.007	4.809	NR	0.010	5.482	NR
Variance-Covariance	0.050	19.601	R	0.027	31.344	R	0.007	6.623	R	0.021	12.178	R	0.022	14.986	R	0.024	13.603	R
Extreme Value Theory	0.009	3.453	NR	0.002	2.038	NR	0.006	5.036	R	0.005	2.872	NR	0.005	3.528	NR	0.005	2.859	NR
AR-ARCH(1)-N	0.026	10.206	R	0.005	5.247	NR	0.009	8.358	R	0.019	10.710	R	0.013	9.009	R	0.012	6.801	R
AR-ARCH(1)-t	0.011	4.266	NR	0.003	3.516	NR	0.004	3.998	NR	0.008	4.536	NR	0.008	5.742	NR	0.006	3.527	NR
AR-GARCH(1.1)-N	0.011	4.259	NR	0.002	2.644	NR	0.007	6.765	R	0.017	9.810	R	0.008	5.481	NR	0.011	5.970	R
AR-GARCH(1.1)-t	0.006	2.497	NR	0.002	1.946	NR	0.006	5.772	R	0.005	3.094	NR	0.005	3.689	NR	0.006	3.660	NR
AR-EGARCH(1.1)-N	0.011	4.486	NR	0.002	2.530	NR	0.009	8.565	R	0.013	7.476	R	0.009	6.201	R	0.011	6.022	R
AR-EGARCH(1.1)-t	0.007	2.593	NR	0.002	1.905	NR	0.006	5.392	R	0.005	2.935	NR	0.006	4.420	NR	0.005	2.828	NR
Monte Carlo Simulation	0.006	2.446	NR	0.002	1.932	NR	0.006	5.808	R	0.006	3.401	NR	0.006	4.418	NR	0.006	3.522	NR
Combination Mean	0.008	3.104	NR	0.004	4.426	NR	0.005	4.452	NR	0.006	3.411	NR	0.007	4.936	NR	0.007	4.232	NR
Combination Median	0.008	3.147	NR	0.002	2.096	NR	0.005	4.849	NR	0.006	3.553	NR	0.007	4.735	NR	0.008	4.466	NR
Combination Trim1	0.006	2.333	NR	0.003	3.376	NR	0.005	4.403	NR	0.005	3.135	NR	0.005	3.739	NR	0.007	3.948	NR
Combination Trim2	0.008	3.172	NR	0.003	3.745	NR	0.005	4.477	NR	0.006	3.460	NR	0.007	4.939	NR	0.007	4.197	NR

Notes: The Table reports the Performance Evaluation results for all the implemented methods. The PM column provides the Penalization Measure while the Ratio column provides the corresponding ratio. R (NR) suggests rejection (non-rejection) with respect to the performance equality test.

Table B12 Performance Evaluation: Second Stage

Panel A: 5%	4 TC CAPE			4 TC PAN			TD3			TD5			TD7			TD9		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.046	7.700	R	0.014	6.869	NR	0.033	6.670	NR	0.053	6.348	R	0.056	8.067	R	0.059	6.809	NR
Historical Simulation-250	0.054	8.895	R	-	-	-	0.040	8.019	R	0.062	7.386	R	0.062	8.972	R	0.072	8.422	R
Historical Simulation-500	0.068	11.343	R	-	-	-	-	-	-	0.064	7.632	R	0.056	8.085	R	0.078	9.052	R
Historical SimulationAll	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Riskmetrics	0.048	7.953	NR	0.021	10.646	R	0.038	7.605	R	0.052	6.253	NR	0.038	5.455	NR	0.044	5.062	NR
Variance-Covariance	-	-	-	-	-	-	-	-	-	0.066	7.895	R	-	-	-	0.075	8.721	R
Extreme Value Theory	-	-	-	0.018	9.097	R	-	-	-	-	-	-	-	-	-	-	-	-
AR-ARCH(1)-N	-	-	-	0.018	9.238	R	0.036	7.117	NR	0.063	7.557	R	0.046	6.682	NR	0.057	6.611	NR
AR-ARCH(1)-t	-	-	-	-	-	-	0.039	7.796	R	-	-	-	-	-	-	-	-	-
AR-GARCH(1.1)-N	0.043	7.072	NR	0.013	6.365	NR	0.035	7.000	NR	0.060	7.197	NR	0.039	5.704	NR	0.051	5.963	NR
AR-GARCH(1.1)-t	0.044	7.230	NR	0.013	6.727	NR	0.039	7.830	R	-	-	-	0.057	8.305	R	0.077	8.991	R
AR-EGARCH(1.1)-N	0.042	7.009	NR	0.012	6.211	NR	0.034	6.823	NR	0.056	6.751	NR	0.040	5.851	NR	0.050	5.794	NR
AR-EGARCH(1.1)-t	0.045	7.533	NR	0.012	6.245	NR	0.035	6.891	NR	0.055	6.637	R	0.056	8.148	R	0.074	8.634	R
Monte Carlo Simulation	0.044	7.323	NR	0.013	6.656	NR	0.037	7.346	R	0.066	7.941	R	0.055	7.971	R	-	-	-
Combination Mean	0.041	6.887	NR	0.018	8.953	R	0.034	6.725	NR	0.059	7.103	NR	0.046	6.650	NR	0.055	6.418	NR
Combination Median	0.043	7.128	NR	0.012	6.222	NR	0.034	6.811	NR	0.056	6.755	NR	0.044	6.304	NR	0.054	6.327	NR
Combination Trim1	0.042	6.907	NR	0.017	8.625	R	0.034	6.701	NR	0.061	7.258	R	0.048	6.902	NR	0.056	6.489	NR
Combination Trim2	0.042	7.020	NR	0.016	8.145	R	0.033	6.666	NR	0.061	7.286	R	0.048	6.904	NR	0.058	6.707	NR
Panel B: 1%	4 TC CAPE			4 TC PAN			TD3			TD5			TD7			TD9		
	PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio		PM	Ratio	
Filtered Historical Simulation	0.007	5.305	NR	0.002	4.223	NR	0.004	5.666	NR	0.006	5.438	NR	0.005	5.379	NR	0.006	5.970	NR
Historical Simulation-250	0.015	11.840	R	0.005	11.956	R	0.006	7.745	R	0.012	11.747	R	-	-	-	0.010	9.018	R
Historical Simulation-500	0.012	9.347	R	-	-	-	0.006	7.551	R	0.012	11.778	NR	0.006	6.393	NR	-	-	-
Historical SimulationAll	-	-	-	-	-	-	0.005	7.053	NR	0.012	11.136	NR	0.008	8.441	R	-	-	-
Riskmetrics	-	-	-	0.004	8.657	R	0.007	8.576	R	0.010	9.093	R	0.007	7.608	R	0.010	9.189	R
Variance-Covariance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Extreme Value Theory	0.009	7.098	R	0.002	4.328	NR	0.006	7.227	R	0.005	4.801	NR	0.005	5.580	NR	0.005	4.793	NR
AR-ARCH(1)-N	-	-	-	0.005	11.139	R	-	-	-	-	-	-	-	-	-	-	-	-
AR-ARCH(1)-t	0.011	8.770	R	0.003	7.466	R	0.004	5.737	NR	0.008	7.582	R	0.008	9.084	R	0.006	5.913	NR
AR-GARCH(1.1)-N	0.011	8.755	R	0.002	5.614	NR	-	-	-	-	-	-	0.008	8.671	R	0.011	10.008	R
AR-GARCH(1.1)-t	0.006	5.134	NR	0.002	4.133	NR	0.006	8.282	R	0.005	5.171	NR	0.005	5.836	NR	0.006	6.136	NR
AR-EGARCH(1.1)-N	0.011	9.223	R	0.002	5.371	NR	-	-	-	-	-	-	-	-	-	0.011	10.095	R
AR-EGARCH(1.1)-t	0.007	5.331	NR	0.002	4.045	NR	0.006	7.737	R	0.005	4.906	NR	0.006	6.993	R	0.005	4.740	NR
Monte Carlo Simulation	0.006	5.028	NR	0.002	4.102	NR	0.006	8.335	R	0.006	5.685	NR	0.006	6.990	NR	0.006	5.904	NR
Combination Mean	0.008	6.381	NR	0.004	9.397	R	0.005	6.388	NR	0.006	5.701	NR	0.007	7.808	R	0.007	7.094	NR
Combination Median	0.008	6.469	NR	0.002	4.450	NR	0.005	6.958	NR	0.006	5.939	NR	0.007	7.490	R	0.008	7.486	R
Combination Trim1	0.006	4.797	NR	0.003	7.168	R	0.005	6.319	NR	0.005	5.239	NR	0.005	5.915	NR	0.007	6.617	NR
Combination Trim2	0.008	6.521	NR	0.003	7.952	R	0.005	6.425	NR	0.006	5.784	NR	0.007	7.813	R	0.007	7.036	NR

Notes: The Table reports the second stage Performance Evaluation results. For each index the worst performing methods are excluded and the analysis is repeated with the remaining methods. The PM column provides the Penalization Measure while the Ratio column provides the corresponding ratio. *R* (*NR*) suggests rejection (non-rejection) with respect to the performance equality test.